# Functional Analysis

## Foundations and Advanced Topics

Martín Argerami

**Martín Argerami**

# Functional Analysis
## Foundations and Advanced Topics

First Edition
*Version 1.11– 2026-09-11*

Martín Argerami
Department of Mathematics and Statistics
University of Regina
Regina, SK, Canada

ISBN 978-1-0698019-0-6 (eBook)

Mathematics Subject Classification: 46-01, 28-01, 47-01, 46A03, 46A22, 46B03, 46C05, 46E30, 46L05, 46L06, 46L07, 46L10

*First Edition. Version 1.11 – 2026-09-11*
This digital version may be updated periodically; the most recent version is available at https://book.argerami.ca/

*To Romina, who has been giving me happiness for decades. And to Jose, Euge, Vicky, Tere, Tom, Santi, Lucas, and Emilia, who have also been doing a pretty good job at it.*

# Preface

These notes were initially developed for the Functional Analysis course that I taught in the Fall of 2010, 2012, 2018, 2020, 2022, and Winter 2025 at the University of Regina. Over the years, in dealing with prerequisites and improving the self-containment of the text, the notes grew to include enough material for a course in Measure Theory (Chapter 2) that I taught several times at the University of Regina, particularly with these notes on Winter 2020, 2021, 2023, and Fall 2024. Eventually I expanded with basic bits of Complex Analysis and topics of interest that go beyond usual material in Functional Analysis and more into Operator Theory and Operator Algebras. Some material and part of the initial format were inspired by the classic books by W. Rudin [Rud87] and J. Conway [Con90], while much was developed in-house and/or from other texts. My answers and feedback at Math.StackExchange have been a big source of material and inspiration.

Significant effort has been put into using material from earlier chapters as reference for the more advanced material, and to make the text as self-contained as possible. When I was a student a very long time ago, I remember being repeatedly frustrated by the statement "this result is beyond the scope of this text/course"; I have strived to avoid that as much as possible. It has been very satisfying, when writing the last chapters, to often find that results I needed were available in previous chapters. It is my hope that readers will find this self-containment useful. A substantial prerequisites chapter contains big chunks of needed previous knowledge, but I do not claim that all is there. Said prerequisites chapter was written with the intent to help the reader broaden their view on already familiar material.

This text has been mostly conceived as an electronic resource giving the reader the advantage of using links within the text, together with the ability to search. Because of that I plan to have regular updates when typos and other feedback come in, so the minor version number will change accordingly.

Some of the graphs that appear in the text have been prepared using *Desmos Graphing Calculator*, under a Creative Commons "Attribution-ShareAlike" License (CC-BY-SA-4.0)

I want to thank Ruy Exel, Ryszard Szwarc, Bruce Blackadar, and Bill Johnson for insightful comments regarding the material; Ken Davidson and Matt Kennedy for generous and useful advice; Shane Crerar, Sushil Singla, and Luis Islas Vizcarra for reading earlier versions and reporting typos; my daughter Eugenia for designing the cover, editorial comments, and logistical help; and I especially thank Doug Farenick and Remus Floricel for the constant support and encouragement.

Martín Argerami
Regina, SK, Canada
November 2025

*This is a good place to mention that I will wholeheartedly welcome any feedback, from typos to general comments to conceptual views. I can be contacted at* `argerami@uregina.ca`.

# Pedagogical Considerations

There is no uniform criterion on whether to take Hilbert spaces or Banach spaces as the starting point for Functional Analysis, and I have hesitated between the two approaches over the years. The former is a particular example of the latter, with more structure (the inner product). I have opted to start with Hilbert spaces to exploit the reader's familiarity with the finite-dimensional case and introduce new ideas in such a setting. Banach spaces are then introduced to go fully into the theory of Functional Analysis. For operators, I took the opposite stance: the general case of operators on a Banach space is covered first, and these results are thus available for the particular case of operators on a Hilbert space. The chapters on Measure Theory and Complex Analysis would not normally form part of a course in Functional Analysis—nor would there be time in a regular semester—but some knowledge of that material will be assumed throughout the text.

The book begins with a prerequisites chapter. It was not written with the intent of teaching the material, but rather provide the reader with a measuring rod and, in some cases, a more advanced perspective; it also makes the book mostly self-contained. While a decent amount of familiarity with most topics in this chapter is required at different stages of the book, it should be possible to advance through parts of the text even if not everything in the chapter is mastered.

The contents in Chapter 2 have been used for a one-semester course in Measure Theory. There is maybe more emphasis than usual on $L^p$ spaces, as these are needed as examples of Banach spaces. Lebesgue measure is developed in parallel with general measure theory, as the effort to prove the general results is often the same as what is required for the particular case of the Lebesgue measure.

The chapter on Complex Analysis was not written with a course in mind. The aim was to be a reference since some results are used in a few crucial points in the book. A secondary reason is that I am fascinated with the Residue Theorem and its applications; plus, it is actually needed for the proof of Lidskii's Theorem in Chapter C.

My basic one-semester Functional Analysis course usually consists of Chapters 4 to 7 and possibly Sections 9.1, 9.2, 9.5 and 9.6. Sections 7.4 and 7.5 could be considered optional, but I do include them in my own basic courses. Over the years, the possibility of covering all the desired material has varied greatly depending on the students' background.

Section 5.6 can be considered as reference. At the same time, several technical details about $L^p$ and $\ell^p$ spaces are provided that are used in later chapters, and other examples are also referenced later. In any case, this section was written with the expectation that getting to see concrete examples of duals and duality could help in understanding the theoretical aspects.

Section 5.7 and Chapter 6 contain the deep core theorems that should appear in any course in Functional Analysis. Similarly mandatory as core topics are the basics of the weak topologies, Sections 7.1 and 7.2.

Sections 7.3 to 7.6 and Chapter 8 develop, in some detail, topics that could be considered optional and can be tailored to the instructor's wishes. Section 7.4 is fairly important in the sense that many of the ideas and results there pave the way for results later in the book. Chapter 9 contains the basics about bounded linear operators over a Banach space. Sections 9.8 and 9.9 showcase some of the amazing sophistication achieved in the study of Banach spaces.

The book then continues with a decent amount of material about Operator Theory and Operator Algebras: bounded operators on a Hilbert space (Chapter 10); C*-algebras (Chapter 11); the Spectral Theorem and basics of von Neumann algebras (Chapter 12); completely positive maps, crossed products, and tensor products (Chapter 13); and von Neumann algebras (Chapter 14). This material could be arranged in two advanced courses, one more introductory with material from Chapters 10 and 11 and a more advanced one with the rest of Chapter 11 and most of Chapters 12 to 14.

The appendices aim to expose some theory and techniques that are used and/or are related to the text, but whose inclusion within the text would impede the flow of the exposition. Some of them, like the detailed proofs of Banach–Tarski and Lidskii's Theorem, do not seem to be widely available—at least with enough detail in the cases known to me—so I had the motivation to fill the gaps.

Most sections end with a list of exercises. The majority of the exercises are included with the aim of improving the student's comprehension of the text, and in many cases require the student to complete the proofs in the exposition. When exercises are quoted in the text, the expectation is that the reader will tackle them right away, as they complement the text and help understanding. Effective study at this level should require the student to attack a large percentage of these exercises. An answer manual is therefore provided with answers to all exercises. The existence of this manual is tricky, in the sense that it can clearly affect learning if consulted too early. Being stuck and stumped at questions is an essential part of learning advanced mathematics, so having the answers at hand could have a detrimental effect. At the same time, answers for many of the questions can be found online with ease, and in many cases the answers found online will actually be my own. So I felt it better to have my answers concentrated and available, at the cost of putting more responsibility on the reader to work on the material and delay looking at the answers. This also allowed me to constrain a few technical aspects of certain proofs to the exercises, where if needed the reader can consult my answer right away.

*Here I reiterate that I will wholeheartedly welcome any feedback, from typos to general comments to conceptual views, and I can be contacted at* `argerami@uregina.ca`.

# Contents

CHAPTER 1

# Prerequisites

We cover here several basic topics that can be considered as prerequisites. This chapter is intended as reference; it should be seen as a refresher, and not as a source to learn the material. At the same time some points of view may help achieve a more sophisticated view of already known ideas.

## 1.1. Set Theory

**1.1.1. Sets and Operations.** As usual, we mostly denote sets with capital letters, and their elements with lowercase letters. This, of course, in the cases where the set is seen as a set, and not as something else; we do not, for instance, see functions as sets, although we define them as such. So, $a \in A$ means that $a$ is an element of the set $A$, and $B \subset A$ means that every element of $B$ is an element of $A$; we say that $B$ is a **subset** of $A$ (or, is **included** in $A$). The sets $A$ and $B$ are equal if they have precisely the same elements, or equivalently if $A \subset B$ and $B \subset A$. The empty set is denoted by $\varnothing$, and the usual set operations are defined as

$$A \cup B = \{c : \ c \in A \text{ or } c \in B\}, \qquad A \cap B = \{c : c \in A \text{ and } c \in B\}.$$

More generally, if $\mathcal{F}$ is a family of sets, then

$$\bigcup_{A \in \mathcal{F}} A = \{a : \ \exists A_0 \in \mathcal{F} \ \text{ with } a \in A_0\}, \qquad \bigcap_{A \in \mathcal{F}} A = \{a : \ \forall A \in \mathcal{F}, \ a \in A\}.$$

The **power set** of $A$ is the set $\mathcal{P}(A)$ of all subsets of $A$. Given a family $\mathcal{F}$ of sets, it is conceptually convenient to consider $J = \mathcal{F}$ as an "index set" and then write the family as $\{A_j\}_{j\in J}$. Looks better than writing $\{A_B\}_{B\in\mathcal{F}}$ and $A_B = B$ (though it means the same!).

So usually a family of sets will be written as $\{A_j\}_{j\in J}$ and

$$\bigcup_{j\in J} A_j = \{a:\ \exists j \in J,\ a \in A_j\}, \qquad \bigcap_{j\in J} A_j = \{a:\ \forall j \in J, a \in A_j\}.$$

We use **family** and **collection** interchangeably, to refer to a $\{A_j\}$.

Usually sets are considered in some **universe** $U$. If $A \subset U$, then the **complement** of $A$ is the set

$$A^c = \{b \in U:\ b \notin A\}.$$

Given two sets $A, B$, the **set difference** of $A$ and $B$ is the set $A \setminus B = A \cap B^c$. When $A \cap B = \varnothing$, we say that $A$ and $B$ are **disjoint**.

The **Cartesian product** of $A$ and $B$ is the set of ordered pairs

$$A \times B = \{(a, b):\ a \in A,\ b \in B\}.$$

The Cartesian product of an arbitrary family of sets will be defined in Section 1.2.

**1.1.2. Relations and Order.** A **relation** $R$ between two sets $A$ and $B$ is a subset $R \subset A \times B$. We write $aRb$ to mean that $(a, b) \in R$. When $A = B$ we say that $R$ is a relation on $A$. An **order relation** on $A$ is a relation $R$ that is

(1) **reflexive:** $aRa$ for all $a \in A$;
(2) **anti-symmetric:** if $aRb$ and $bRa$, then $a = b$;
(3) **transitive:** if $aRb$ and $bRc$, then $aRc$.

An obvious well-known order relation is the usual order of numbers. This an example of a **total order**: given any $a, b \in \mathbb{R}$, either $a \le b$ or $b \le a$. We do not require this from orders in general. A prototypical order is inclusion of sets; if we denote this order by $\le$, then we have for example that $\{1, 2\} \le \{1, 2, 3\}$ and $\{2, 3\} \le \{1, 2, 3\}$; but $\{1, 2\}$ and $\{2, 3\}$ are not related. Reversing an order relation is also an order relation; a typical situation where this is used is when we consider reverse inclusion of sets as the order.

Given an order relation $\le$ on a set $A$ and $B \subset A$, an **upper bound** for $B$ is any element $a \in A$ such that $b \le a$ for all $b \in B$. A **lower bound** for $B$ is an $a \in A$ such that $a \le b$ for all $b \in B$. A set is **bounded above** when it admits an upper bound; and, similarly, **bounded below** when it admits a lower bound. The ordered set $(A, \le)$ is **complete** (or, we say that $A$ is **completely ordered**) if the order is total and if every bounded above set admits a least upper bound. The **least upper bound** of a subset $B \subset A$ is the **supremum** of $B$, denoted $\sup B$. When there is a **greatest lower bound** for $B$, we call it the **infimum** of $B$ and denote it by $\inf B$.

**1.1.3. Functions.** Given sets $A, B$, a **function** $f$ with domain $A$ and codomain $B$ is a relation $f \subset A \times B$ such that for each $a \in A$ there exists a unique $b \in B$ with $(a, b) \in f$. It is also common to use the words **map** and **mapping** for a function. The usual notation is, when the pair $(a, b)$ is in $f$, to write $f(a) = b$; we also use the notation $a \longmapsto f(a)$. The function is commonly denoted by $f : A \to B$ and the set $f \subset A \times B$ is most often referred as the **graph** of $f$.

Given a function $f : A \to B$, the **image** of $A$ is the set $f(A) = \{f(a) : \ a \in A\} \subset B$. For a subset $B_0 \subset B$, the **preimage** of $B_0$ by $f$ is the set

$$f^{-1}(B_0) = \{a \in A : \ f(a) \in B_0\}.$$

If $f : A \to B$ and $g : B \to C$ are functions, the **composition** $g \circ f$ is the function $g \circ f : A \to C$ given by $g \circ f(a) = g(f(a))$. The **identity function** is the function $\mathrm{id}_A : A \to A$ given by $\mathrm{id}_A(a) = a$ for all $a \in A$.

A very useful—and repeatedly used—fact is that preimages behave perfectly with respect to set operations:

***1.1.1. Proposition.***

*Let $f : A \to B$ be a function, $\{A_j\}_{j \in J}$ be a collection of subsets of $A$, and $\{B_k\}_{k \in K}$ a collection of subsets of $B$. Then*

$$f^{-1}\Big(\bigcup_k B_k\Big) = \bigcup_k f^{-1}(B_k); \qquad f^{-1}\Big(\bigcap_k B_k\Big) = \bigcap_k f^{-1}(B_k);$$

*and*

$$f^{-1}(B \setminus B_0) = A \setminus f^{-1}(B_0).$$

*Also,*

$$f\Big(\bigcup_j A_j\Big) = \bigcup_j f(A_j), \qquad f\Big(\bigcap_j A_j\Big) \subset \bigcap_j f(A_j)$$

*This last inclusion can be strict.*

*Proof.* Exercise 1.1.4. □

A function $f : A \to B$ is

- **injective**, or **one-to-one**, if $f(a_1) = f(a_2)$ implies $a_1 = a_2$;
- **surjective**, or **onto**, if $f(A) = B$;
- **bijective** if it is both injective and surjective.

The function $f$ is bijective if and only if there exists $g : B \to A$ with $g(f(a)) = a$ for all $a \in A$, and $f(g(b)) = b$ for all $b \in B$; that is, $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$ (Exercise 1.1.6). In that case we say that $f$ is **invertible** and the function $g$ is called the **inverse** of $f$; and, since when it exists it is unique (Exercise 1.1.7), we denote it by $f^{-1}$.

The set of all functions $f : A \to B$ is often denoted $B^A$.

**1.1.4. Equivalence Relations.** An **equivalence relation** on $A$ is a relation $R$ on $A$ that is

(1) **reflexive:** $aRa$ for all $a \in A$;
(2) **symmetric:** if $aRb$ then $bRa$;
(3) **transitive:** if $aRb$ and $bRc$, then $aRc$.

The point of an equivalence relation is that it somehow encapsulates the properties of equality. The **equivalence class** of $a$ is the set

$$[a] = \{b \in A : \ aRb\}.$$

It is straightforward to show that $aRb$ if and only if $[a] = [b]$. As a consequence, the equivalence classes of $A$ under $R$ form a **partition**: two equivalence classes are either equal or disjoint, and every element of $A$ belongs to precisely one equivalence class. The **quotient** of $A$ by $R$ is the set $A/R$ of equivalence classes of $R$. A **representative** of the class $[a]$ is any element $b \in [a]$.

Often the set $A$ has some additional structure and/or operations, and one might want to use them to define them on the equivalence classes. This usually requires one to check that the operation is **well-defined**, in the sense that the operation on classes is defined in terms of representatives, and it is necessary to confirm that the definition does not depend on the chosen representatives. To mention an example that is particularly relevant to this text, let $V$ be a vector space over $\mathbb{C}$ and $W \subset V$ a subspace. Then the quotient space $V/W$ is the set $V/\sim_W$, where $v_1 \sim_W v_2$ if $v_1 - v_2 \in W$; one should know—or otherwise check—that this is an equivalence relation, and the classes are the sets $v + W$ for $v \in V$. Because $W$ is a subspace, the equivalence relation is compatible with the vector space operations, in the sense that if $v_1 \sim_W v_1'$, $v_2 \sim_W v_2'$, then $av_1 + v_2 \sim_W av_1' + v_2'$. This allows one to define vector space operations in the quotient, namely $a(v_1 + W) + (v_2 + W) = av_1 + v_2 + W$, and $V/W$ becomes a vector space. In the case of an algebra $A$, one needs an ideal $I \subset A$ to be able to define an algebra structure on $A/I$.

Further examples of quotients can be seen in Section 1.3.

**1.1.5. Exercises.**

(1.1.1) Let $\{A_j\}_{j\in J}$ be a collection of subsets of a set $A$. Show that

$$\Big[\bigcup_{j\in J} A_j\Big]^c = \bigcap_{j\in J} A_j^c, \qquad \Big[\bigcap_{j\in J} A_j\Big]^c = \bigcup_{j\in J} A_j^c$$

(1.1.2) For sets $A, B, C$, show that

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

and

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

(1.1.3) Let $\{A_j\}_{j\in J}, \{B_j\}_{j\in J}$ be collections of subsets of a set $A$. Are the equalities

$$\Big[\bigcup_{j\in J} A_j\Big] \cap \Big[\bigcup_{j\in J} B_j\Big] = \bigcup_{j\in J}(A_j \cap B_j)$$

and

$$\Big[\bigcap_{j\in J} A_j\Big] \cup \Big[\bigcap_{j\in J} B_j\Big] = \bigcap_{j\in J}(A_j \cup B_j)$$

true? Prove them, or find a counterexample.

(1.1.4) Prove Proposition 1.1.1. Show that the inclusion $f\Big(\bigcap_j B_j\Big) \subset \bigcap_j f(B_j)$ can be strict.

(1.1.5) Let $f : A \to B$. For any $B_0 \subset B$ and $A_0 \subset A$, show that

$$f(f^{-1}(B_0)) \subset B_0, \qquad A_0 \subset f^{-1}(f(A_0)).$$

Show that equality does not always hold, but that

$$f(f^{-1}(B_0)) = B_0$$

whenever $f$ is surjective or, more generally, if $B_0 \subset f(A)$.

(1.1.6) Let $f : A \to B$ be a function. Show that

(i) $f$ is injective if and only if there exists $g : B \to A$ with $g \circ f = \mathrm{id}_A$;

(ii) $f$ is surjective if and only if there exists $h : B \to A$ with $f \circ h = \mathrm{id}_B$;

(iii) $f$ is bijective if and only if it is invertible.

(1.1.7) Let $A$ be a set with an associative operation $(a, b) \longmapsto ab$ and with a unit $e \in A$ (that is, $ae = ea = a$ for all $a \in A$). Show that the unit is unique. Show also that if $a \in A$ is invertible (that is there exists $b \in A$ with $ab = ba = e$) then $b$ is unique with that property. More generally, show that if $a$ has a left inverse $b$ and a right inverse $c$, then $b = c$.

(1.1.8) Let $R$ be a relation on $\mathbb{Z}$ defined by:

$$a\,R\,b \iff 3 \text{ divides } (a-b).$$

Show that $R$ is an equivalence relation. Determine its equivalence classes. The quotient $\mathbb{Z}/R$ is often denoted by $\mathbb{Z}_3$.

(1.1.9) Let $f : X \to Y$ be a function. Define a relation $\sim$ on $X$ by:

$$x_1 \sim x_2 \iff f(x_1) = f(x_2).$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that the equivalence classes are the "fibers" of $f$ (i.e., sets of the form $f^{-1}(\{y\})$ for $y \in Y$).

(1.1.10) Consider the relation $\sim$ on $\mathbb{R}^2$ defined by:

$$(x_1, y_1) \sim (x_2, y_2) \iff x_1^2 + y_1^2 = x_2^2 + y_2^2.$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Describe the equivalence classes geometrically.

(1.1.11) Let $S = \{1, 2, 3, 4, 5\}$ and define $R$ as:

$$R = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (1, 2), (2, 1), (3, 4), (4, 3)\}.$$

(i) Show that $R$ is an equivalence relation.

(ii) Determine the quotient $S/R$.

(1.1.12) Prove that any partition $\mathcal{P}$ of a set $S$ induces an equivalence relation $\sim$ on $S$ where:

$$a \sim b \iff a \text{ and } b \text{ belong to the same subset in } \mathcal{P}.$$

Conversely, show that any equivalence relation on $S$ induces a partition of $S$.

(1.1.13) Let $\sim$ be an equivalence relation on $\mathbb{N} \times \mathbb{N}$ defined by:

$$(a, b) \sim (c, d) \iff a + d = b + c.$$

*Attention: the set $R$ here consists of pairs of ordered pairs!*

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that addition defined as:

$$[(a, b)] + [(c, d)] = [(a + c, b + d)]$$

is **well-defined** (i.e., independent of the choice of representatives).

(iii) Show that multiplication defined as

$$[(a, b)][(c, d)] = [ac + bd, ad + bc]$$

is well-defined, and it is distributive with respect to addition.

(1.1.14) Let $\sim$ be an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$ (where $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$) defined by:

$$(a, b) \sim (c, d) \iff ad = bc.$$

(i) Prove that $\sim$ is an equivalence relation.

(ii) Show that addition defined as:

$$[(a, b)] + [(c, d)] = [(ad + bc, bd)]$$

is **well-defined** (i.e., independent of the choice of representatives).

(iii) Show that multiplication defined as

$$[(a, b)][(c, d)] = [ab, cd]$$

is well-defined, and it is distributive with respect to addition.

## 1.2. The Axiom of Choice

The most common framework for set theory—and all of mathematics—is the so-called **Zermelo–Fraenkel** theory (**ZF**). These are a basic list of axioms that provide a decent start to develop set theory as we understand it. We will not discuss these concrete axioms here, but they allow the manipulations in the previous section. The axioms in ZF also provide the existence of the natural numbers (more properly, the von Neumann ordinal $\omega$) but are kind of lacking in tools to address infinite sets. So it is more or less clear some additional axiom is needed.

Having functions available, we can make the following definition. The (Cartesian) **product** of the family $\{A_j\}_{j\in J}$ is the set

$$\prod_{j\in J} A_j = \{f : J \to \bigcup_{j\in J} A_j \text{ such that } \; f(j) \in A_j\}.$$

This definition of product is sometimes hard to grasp. Starting with the familiar case of sequences and their notation $\{a_n\}$, what may help is to notice that if instead of writing $a_n$ we write $a(n)$, it is clear that a sequence of real numbers, say, is a function $a : \mathbb{N} \to \mathbb{R}$. So for instance the set of all real-valued sequences is $\prod_{n\in\mathbb{N}} \mathbb{R}$, often denoted $\mathbb{R}^{\mathbb{N}}$. Because of this it is common, as has been mentioned before, to denote the set of all functions $A \to B$ as $B^A$.

It is standard in Functional Analysis to assume the Axiom of Choice. This has deep consequences, though, so it is not taken lightly; see Chapter D. Consciously or not, the Axiom of Choice was used in the answer to Exercise 1.1.6.

We will discuss these axiomatic frameworks a bit more on Chapter D, where a very non-intuitive consequence of the Axiom of Choice, the Banach–Tarski Paradox, is considered.

***The Axiom of Choice.*** *Given a set $J$ and a family of nonempty sets $\{A_j\}_{j\in J}$, the Cartesian product $\prod_{j\in J} A_j$ is nonempty.*

In a naive view, it is hard to imagine what could possibly be problematic about the Axiom of Choice. How could there not exist a function $f : J \to \bigcup_j A_j$ such that $f(j) \in A_j$ for all $j$ (a "choice")? Let us consider first a couple more (possible) properties of sets.

***The Well-Ordering Principle.*** *Every set admits a well-order. That is, on any set there exists a total order such that every non-empty subset has a least element.*

The typical example of a well-order is that of the natural numbers. The fact that the natural numbers are well-ordered is equivalent with the Induction Principle, and every construction of the natural numbers is crafted so that these properties hold. In general, though, even for totally ordered sets like $\mathbb{R}$, one should not expect the well-ordering to come from the original order of the set.

Given an ordered set $Z$, a **chain** in $Z$ is a subset $C \subset Z$ that is totally ordered.

***Zorn's Lemma.*** *Let $Z$ be a partially ordered set. If every chain in $Z$ admits an upper bound in $Z$, then $Z$ has a maximal element.*

There is a folklore joke among mathematicians that says that the Axiom of Choice is obviously true, the Well-Ordering Principle is obviously false, and who knows about Zorn's Lemma. The joke is funny because it turns out that these three axioms are equivalent.

A common mindset is then to assume the Axiom of Choice, and thus enjoy the advantage of having Zorn's Lemma at our disposal. This is often called the Zermelo–Fraenkel + Choice (**ZFC**) framework, and it is the setup we adopt; it is the standard in Measure Theory and in Functional Analysis. Zorn's Lemma is crucial to prove the existence of many different kinds of objects when dealing with infinite dimension.

Let us check a classic example of how to use Zorn's Lemma, by showing that any vector space $V$ has a basis. Let $\mathcal{B}$ be the set of linearly independent families $\{v_j\} \subset V$. The set $\mathcal{B}$ is nonempty as it contains the singletons $\{v\}$ for each nonzero $v \in V$. We make $\mathcal{B}$ an ordered set by using inclusion as our order. Let $\{B_k\} \subset B$ be a chain (the indices do not have to be integers, they could be any totally ordered set). This means that each $B_k$ is a family of linearly independent vectors, the indices are all comparable, and $B_j \subset B_k$ if $j \leq k$. A routine computation shows that $\bigcup_k B_k$ is in $\mathcal{B}$ (i.e., its elements are linearly independent) and it is an upper bound for $\{B_k\}$. By Zorn's Lemma, $\mathcal{B}$ admits a maximal element, which will be a maximal linearly independent set in $V$: a basis. If this seems unnecessarily complicated, let us remark that models of ZF (without Choice) exist where one can find a vector space without any basis.

**1.2.1. Exercises.** *(These exercises could be tricky depending on background. The reader who is in doubt should just take a look at the book of answers.)*

(1.2.1) Let $R$ be a nonzero unital commutative ring and $x \in R$ non-invertible. Show that there exists a proper maximal ideal $J$ of $R$ with $x \in J$.

(1.2.2) Let $X$ be a set. Prove that Zorn's Lemma implies the Well Ordering Principle by applying Zorn's Lemma to the collection

$$\mathcal{P} = \{(A, \leq_A) : A \subset X,\ \leq_A \quad \text{is a well-ordering on } A\},$$

where the order is given by saying that $(A, \leq_A) \preceq (B, \leq_B)$ if $A \subset B$, $\leq_B$ extends $\leq_A$, and every element of $A$ is less (in the $\leq_B$ order) than every element of $B \setminus A$.

## 1.3. Real Numbers and Calculus

Notation for the usual sets of numbers are $\mathbb{N} = \{1, 2, 3, \ldots\}$ for the natural numbers, $\mathbb{Z}$ for the integers, $\mathbb{Q}$ for the rationals, and $\mathbb{R}$ for the reals. These sets can be seen to arise from the need to solve certain kinds of equations, and can be constructed from $\mathbb{N}$ in the following usual way:

- We have the problem that equations like $x + 1 = 0$ have no solution on $\mathbb{N}$; or, more abstractly, there is no neutral element nor inverses for addition. We construct $\mathbb{Z}$ as the quotient of $\mathbb{N} \times \mathbb{N}$ by the equivalence relation $(a, b) \sim (c, d)$ if $a + d = b + c$ (this is known as the **Grothendieck group of a commutative monoid**). Define addition by $[(a, b)] + [(c, d)] = [(a + c, b + d)]$. Since this new operation is defined in terms of representatives, we need to check that it is well-defined (Exercise 1.1.13). The natural numbers $\mathbb{N}$ embed in $\mathbb{Z}$ by $n \longmapsto [(n, 0)]$; the additive inverse of $[(n, 0)]$ is $[(0, n)]$; that is, $[(0, n)]$ for $n \in \mathbb{N}$ is the object we usually denote by $-n$, and $[(7, 7)]$ (any other number works the same) is the additive identity, 0.
- Having the integers, we find ourselves unable to solve equations like $2x - 1 = 0$; or more abstractly, $\mathbb{Z}$ is a commutative ring but not a field, as most nonzero elements have no multiplicative inverses. We construct $\mathbb{Q}$ as the quotient of $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ by the equivalence relation $(a, b) \sim (c, d)$ if $ad = bc$. Define addition and multiplication by
$$[(a, b)] + [(c, d)] = [(ad + bc, bd)], \qquad [(a, b)][(c, d)] = [(ac, bd)].$$
Once again it is necessary to check that these two operations are well-defined, and that they satisfy the usual properties we expect from them; that is all routine work that was included in Exercise 1.1.14. The (class of the) pair $[(a, b)]$ is what we usually think of as $\frac{a}{b}$.
- At this stage, $\mathbb{Q}$ is a field and so we can solve any equation of the form $ax + b = 0$ for $a, b \in \mathbb{Z}$. But quadratic equations like $x^2 - 2 = 0$ have no solution in $\mathbb{Q}$. More abstractly, we also have the problem that $\mathbb{Q}$ is not

complete with its usual topology (nor in order: $\mathbb{Q}$ has bounded above sets that admit no least upper bound). We construct $\mathbb{R}$ as the set of Dedekind cuts of $\mathbb{Q}$. A **cut** is a proper subset $E \subset \mathbb{Q}$ with no maximum element, and such that $a \in E$, $b < a$, imply that $b \in E$. Intuitively, "stop somewhere in the real line, and take all numbers to the left of that point". It is also possible to define $\mathbb{R}$ as the quotient of the set of Cauchy sequences in $\mathbb{Q}$, under the equivalence relation $\{x_n\} \sim \{y_n\}$ if $\lim_n |x_n - y_n| = 0$. In both cases, one needs to sit down and carefully define addition, multiplication, order, and check that all properties of a completely ordered field are satisfied.

The actual construction used to obtain $\mathbb{R}$ does not matter much, as $\mathbb{R}$ can be characterized as the unique completely ordered field; this means to say that $\mathbb{R}$ is a field, with a total order, a positive cone $\mathbb{R}^+$ that is closed under addition and multiplication, with the property any for any nonzero $x \in \mathbb{R}$ either $x > 0$ or $-x > 0$, and such that any set bounded above admits a least upper bound (**supremum**). Because in $\mathbb{R}$ a subset $E$ is bounded below if and only if $-E$ is bounded above, the completeness also guarantees the existence (for a bounded below set) of the greatest lower bound (**infimum**). It is fairly straightforward to check that

$$\inf E = -\sup(-E).$$

Although suprema and infima are commonly used in analysis, their definition is purely algebraic; the definition makes sense in any ordered set. Whether suprema and infima exist is a different issue, and it is a wonderful property of the real line that every set bounded above admits a supremum (and an infimum if bounded below). Ultimately, what makes them central in analysis is the fact that the topology in $\mathbb{R}$ is defined via the order. If this sounds like new information, we invite the reader to think about how an open interval in $\mathbb{R}$ is defined. We will go back to this when talking about topology.

When the set $E$ is unbounded above, we write $\sup E = \infty$, and when it is unbounded below we write $\inf E = -\infty$.

Given a sequence $\{a_n\}$ of real numbers we define

$$\limsup_n a_n = \inf_m \sup_{n\geq m} a_n, \qquad \liminf_n a_n = \sup_m \inf_{n\geq m} a_n.$$

The interior sup/inf are monotone, so the outer inf/sup are limits, justifying the notation. It is not hard to show that $\limsup_n a_n$ and $\liminf_n a_n$ are respectively the maximum and minimum cluster points of $\{a_n\}$ (Exercise 1.3.2; for the reader needing a refresher, cluster points are defined on Section 1.8). The analog definition works for functions $f : \mathbb{R} \to \mathbb{R}$. Namely,

$$\limsup_{t\to\infty} f(t) = \inf_s \sup_{t\geq s} f(t), \qquad \liminf_{t\to\infty} f(t) = \sup_s \inf_{t\geq s} f(t).$$

One also considers the limsup and liminf of a real-valued function at a point. Namely,

$$\limsup_{t\to t_0} f(g) = \inf_{\varepsilon>0} \sup_{|t-t_0|<\varepsilon} f(t), \qquad \liminf_{t\to t_0} f(g) = \sup_{\varepsilon>0} \inf_{|t-t_0|<\varepsilon} f(t).$$

We now talk Calculus briefly. A crucial result about differentiation in Calculus is the Mean Value Theorem:

**1.3.1. Theorem. (Mean Value Theorem)**

*Let $f : [a,b] \to \mathbb{R}$ be continuous, and differentiable on $(a,b)$. Then there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)(b-a). \tag{1.1}$$

*Proof.* Assume first that $f(a) = f(b) = 0$ (the result in this setting is known as **Rolle's Theorem**). If $f = 0$ everywhere, the equality is trivial. Otherwise, by Exercise 1.8.37[1] and Heine–Borel (Theorem 1.8.18) $f$ attains its maximum and minimum in $[a,b]$. As $f(a) = f(b) = 0$ and $f \neq 0$, either the maximum or the minimum is not zero. By replacing $f$ with $-f$ if needed, we may assume that there exists $c \in (a,b)$ with $f(c) > 0$ and $f(x) \leq f(c)$ for all $x \in [a,b]$. This means that $f'(c) = 0$ and (1.1) holds.

For arbitrary $f$, we apply the above to

$$g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a) - f(a).$$

So there exists $c \in (a,b)$ such that $g'(c) = 0$. That is,

$$0 = f'(c) - \frac{f(b)-f(a)}{b-a}.$$

□

**1.3.2. Definition.**

*A **partition** of the interval $[a,b]$ is a finite ordered set*

$$P = \{x_0, x_1, \ldots, x_r\} \subset [a,b]$$

*with $x_0 = a$ and $x_r = b$. The supremum and infimum of $f$ over each interval determined by $P$ are denoted*

$$M_j(f,P) = \sup\{f(x) : \ x \in [x_{j-1}, x_j]\}$$

*and*

$$m_j(f,P) = \inf\{f(x) : \ x \in [x_{j-1}, x_j]\}.$$

[1] Removing the incoherence of having a forward reference here was not worth the shuffling of material, as this is a prerequisites chapter

*The **lower and upper sums** of $f$ on $P$ are*

$$L(f,P)=\sum_{j=1}^{r} m_j(f,P)\,(x_j-x_{j-1}),\quad U(f,P)=\sum_{j=1}^{r} M_j(f,P)\,(x_j-x_{j-1}).$$

*A bounded function $f:[a,b]\to\mathbb{R}$ is **Riemann integrable** if*

$$\sup\{L(f,P):\ P\}=\inf\{U(f,P):\ P\}.$$

*In case of equality this number is called the **Riemann integral** of $f$ on $[a,b]$ and denoted by*

$$\int_a^b f(x)\,dx$$

It is easy to show that continuous functions are Riemann integrable, and that even functions with finitely many discontinuities are Riemann integrable. But it is not clear at all what would be a necessary and sufficient condition for Riemann integrability. We will provide such condition in Proposition 2.5.17.

The biggest theorem in calculus is the **Fundamental Theorem of Calculus**, which relates integrals and derivatives.

***1.3.3. Theorem. (Fundamental Theorem of Calculus)***

*Let $f:[a,b]\to\mathbb{R}$ be continuous, and let*

$$F(x)=\int_a^x f(t)\,dt.$$

*Then $F$ is differentiable on $[a,b]$ and $F'(x)=f(x)$.*

*Proof.* Fix $\varepsilon>0$. As $f$ is continuous on the compact set $[a,b]$, it is uniformly continuous. So there exists $\delta>0$ such that $|f(x)-f(y)|<\varepsilon$ whenever $|x-y|<\delta$. We have, if $|h|<\delta$,

$$\begin{aligned}\left|\frac{F(x+h)-F(x)}{h}-f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h}(f(t)-f(x))\,dt\right|\\ &\le \frac{1}{h}\int_x^{x+h}|f(t)-f(x)|\,dt \le \frac{1}{h}\int_x^{x+h}\varepsilon\,dt=\varepsilon.\end{aligned}$$

As this can be done for all $\varepsilon>0$, $F$ is differentiable and $F'(x)=f(x)$. □

One of many useful consequences of the Fundamental Theorem of Calculus is that if $f$ is $n$-times continuously differentiable, then we can write $f$ in terms of its **Taylor Polynomial**:

**1.3.4. Theorem.** *(Taylor Theorem, Lagrange Form)*

*Let $f : [a, b] \to \mathbb{R}$ such that its first $n$ derivatives exist and are continuous on $[a, b]$ and differentiable on $(a, b)$; and such that $f^{n+1}$ exists. Fix $x_0 \in [a, b]$. Then for each $x \in [a, b] \setminus \{x_0\}$ there exists $\xi$ between $x_0$ and $x$ such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{n+1}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \tag{1.2}$$

*Proof.* Note that $x$ and $x_0$ are distinct and fixed. Let

$$M = \frac{-1}{(x - x_0)^{n+1}} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Consider the function $f$ given by

$$F(t) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x - t)^k + M(x - t)^{n+1}.$$

By hypothesis $F$ is continuous on $[a, b]$ and differentiable on $(a, b)$, so the Mean Value Theorem applies: there exists $\xi$ between $x_0$ and $x$ such that

$$F'(\xi) = \frac{F(x) - F(x_0)}{x - x_0} = 0$$

(note that the choice of $M$ guarantees that $F(x_0) = 0$). Now we evaluate $F'(\xi)$. We have

$$0 = F'(\xi) = \sum_{k=0}^{n} \frac{f^{k+1}(\xi)}{k!}(x - \xi)^k - \sum_{k=1}^{n} \frac{f^{(k)}(\xi)}{(k-1)!}(x - \xi)^{k-1} - (n+1)M(x - \xi)^n$$

$$= \frac{f^{n+1}(\xi)}{n!}(x - \xi)^n - (n+1)M(x - \xi)^n$$

.

Therefore $M = f^{n+1}(\xi)/(n+1)!$, and the proof is complete. ☐

Here is an alternative formula for the remainder, easier to obtain via integration by parts:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!}x^k + \frac{x^n}{(n-1)!}\int_0^1 (1 - t)^{n-1} f^{(n)}(tx)\, dt \tag{1.3}$$

(Exercise 1.3.4).

**1.3.5. Definition.**

*Let $f : [a,b] \to \mathbb{R}$ be bounded. The* ***oscillation*** *of $f$ at $x \in [a,b]$ is the number*

$$o(f,x) = \lim_{\delta \to 0} M(f,x,\delta) - m(f,x,\delta),$$

*where*

$$M(f,x,\delta) = \sup\{f(z) : \ z \in (x-\delta, x+\delta)\}$$

*and*

$$m(f,x,\delta) = \inf\{f(z) : \ z \in (x-\delta, x+\delta)\}.$$

A bounded function $f$ is continuous at $x$ if and only if $o(f,x) = 0$ (Exercise 1.3.9). We will need the following result.

***1.3.6. Lemma.***

*Let $f : [a,b] \to \mathbb{R}$ be bounded and $\varepsilon > 0$. Then the set $\{x \in [a,b] : o(f,x) \geq \varepsilon\}$ is compact.*

*Proof.* Write $N = \{x \in [a,b] : \ o(f,x) \geq \varepsilon\}$. As $[a,b]$ is compact, all we need is to show that $N$ is closed. Suppose that $x \in [a,b] \setminus N$. This means that there exists $\delta > 0$ with $M(f,x,\delta) - m(f,x,\delta) < \varepsilon$. Fix $y \in (x-\delta, x+\delta)$ and let $\delta_1 > 0$ such that $(y-\delta_1, y+\delta_1) \subset (x-\delta, x+\delta)$. We have

$$M(f,y,\delta_1) - m(f,y,\delta_1) \leq M(f,x,\delta) - m(f,x,\delta) < \varepsilon.$$

Hence $o(f,y) < \varepsilon$ and therefore $[a,b] \setminus N$ is open. Thus $N$ is closed. □

**1.3.1. Exercises.**

(1.3.1) Let $E \subset \mathbb{R}$. Show that $\inf E = -\sup(-E)$ and

$$\liminf_n a_n = -\limsup_n (-a_n).$$

(1.3.2) Let $\{a_n\} \subset \mathbb{R}$ be a sequence. Allowing $\pm\infty$ to be cluster points for unbounded sequences, show that

$$\limsup_n a_n = \max\{\text{cluster points of } \{a_n\}\},$$

and

$$\liminf_n a_n = \min\{\text{cluster points of } \{a_n\}\}.$$

(1.3.3) Let $\{a_n\} \subset \mathbb{R}$ be a sequence. Show that $\lim_n a_n$ exists and is equal to $L \in \mathbb{R}$ if and only if $\limsup_n a_n = \liminf_n a_n = L$.

(1.3.4) Prove (1.3).

(1.3.5) Let $f : [0,\infty) \to \mathbb{R}$ be continuously differentiable, with $f'(x) > 0$ for all $x$, and such that $f(x) \leq c$ for all $x$.

(i) Show that $\lim_{x\to\infty} f(x)$ exists and it is equal to $\sup\{f(x) : x \geq 0\}$;

(ii) show that there exists $f$ as above and such that $\lim\limits_{x\to\infty} f'(x)$ does not necessarily exist;

(iii) show that if in addition $f'$ is differentiable and $f''(x) < 0$ for all sufficiently large $x$, then $\lim\limits_{x\to\infty} f'(x) = 0$.

(1.3.6) Let $[a,b] \subset \mathbb{R}$ and $P, P'$ partitions with $P \subset P'$. Show that $L(f,P) \leq L(f,P') \leq U(f,P') \leq U(f,P)$.

(1.3.7) Let $f : [a,b] \to \mathbb{R}$ be bounded. Show that $f$ is Riemann integrable if and only if for each $\varepsilon > 0$ there exists a partition $P$ of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$.

(1.3.8) Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Show that

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{k=1}^n f\Big(a + \frac{k(b-a)}{n}\Big)\,\frac{1}{n}. \tag{1.4}$$

(1.3.9) Show that a bounded function $f$ is continuous at $x$ if and only if $o(f,x) = 0$.

## 1.4. Trigonometric Functions

As already mentioned, this chapter is not written as a source to learn the material; the goal is to serve as a recap and a reference, and this is particularly true of this section. While a basic command of trigonometry is assumed, our goal here is to show that trigonometry can be developed analytically.

In Calculus, there is a dichotomy when introducing trigonometric functions. We are told that the sine and the cosine are defined as ratios of sides of square triangles. Some geometry is required to both argue that these ratios do depend on just the angle, and to calculate their derivatives. From there the Taylor series for both functions are derived. This feels a bit cumbersome, and it is not entirely obvious how to make it work formally. We will argue here that the reverse process can be used.

Define

$$s(x) = \sum_{k=0}^\infty (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

This is a power series with infinite radius of convergence in $\mathbb{R}$, which guarantees that we are allowed to differentiate and to take antiderivatives term by term.

Denote its derivative by $c(x) = s'(x)$. A direct computation shows that

$$c(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

and that $c'(x) = -s(x)$. So $c(x)c'(x) + s(x)s'(x) = 0$. Multiplied by 2, this is the derivative of $c(x)^2 + s(x)^2$, which is thus constant; using that $s(0) = 0$ and $c(0) = 1$, we obtain

$$c(x)^2 + s(x)^2 = c(0)^2 + s(0)^2 = 1, \qquad x \in \mathbb{R}. \tag{1.5}$$

One can also get from manipulating the series that

$$s(-x) = -s(x), \qquad c(-x) = c(x),$$

and that

$$\begin{aligned} s(x+y) &= s(x)c(y) + s(y)c(x), \\ c(x+y) &= c(x)c(y) - s(x)s(y). \end{aligned} \tag{1.6}$$

Proving these last two equalities via the series is a bit messy; but it is also possible to deduce (1.6) in a different way. Indeed, fix $y$ and define $f(x) = s(x+y)$. Then

$$f''(x) = -f(x), \qquad f(0) = s(y), \qquad f'(0) = c(y). \tag{1.7}$$

Posing that $f(x) = \sum_{k=0}^{\infty} a_k x^k$, one readily obtains that

$$a_{2k} = \frac{(-1)^k s(y)}{(2k)!}, \qquad a_{2k+1} = \frac{(-1)^k c(y)}{(2k+1)!}.$$

It follows immediately, since the problem (1.7) has unique solution (Lemma 1.4.1 below), that

$$s(x+y) = f(x) = c(x)s(y) + s(x)c(y).$$

And here is the proof that the initial-value-problem (1.7) has a unique solution:

***1.4.1. Lemma.***

*Let $g : [-a, a] \to \mathbb{R}$ be a twice differentiable function such that $g''(x) = -g(x)$, $g(0) = 0$, $g'(0) = 0$. Then $g = 0$. In particular, the IVP $g'' = -g$, $g(0) = a$, $g'(0) = b$, has a unique solution.*

*Proof.* We first note that the condition $g'' = -g$ implies that $g$ is infinitely differentiable. Indeed, by assumption we have that $g$ is twice differentiable, but then $g'' = -g$ is twice differentiable, so $g^{(4)}$ exists, and repeating this we get that $g^{(n)}$ exists for all $n$.

Since $g$ is continuous, there exists $c > 0$ with $|g(x)| \le c$ for all $x \in [-a, a]$. Now fix $x \in [-a, a]$. We have $g''(0) = -g(0) = 0$, and $g'''(0) = -g'(0) = 0$. Iterating this, we get that $g^{(n)}(0) = 0$ for all $n$. Also $|g^{(2n)}(x)| = |g(x)| \le c$.

Now the degree $2n-1$ Taylor approximation for $g$ looks like

$$|g(x)| = \left| g^{(2n)}(\xi)\,\frac{x^{2n}}{(2n)!} \right| \le \frac{cx^{2n}}{(2n)!}.$$

As the inequality holds for all $n$,

$$|g(x)| \le \limsup_{n\to\infty} \frac{cx^{2n}}{(2n)!} = 0. \qquad \square$$

From (1.5) we get that $-1 \le s(x) \le 1$ for all $x$. Suppose that $c(x) > 0$ for all $x$; this would imply that $s(x)$ is always increasing and bounded, so $\lim_{x\to\infty} c(x) = 0$ by Exercise 1.3.5, but this would contradict

$$c(2x) = c(x)^2 - s(x)^2 = 2c(x)^2 - 1.$$

Similarly, we cannot have $c(x) < 0$ for all $x$; so $c(x)$ has at least one zero. Continuity and $c(0) = 1$ guarantee that there exists a least positive $\beta > 0$ with $c(\beta) = 0$. As $s'(0) = c(0) = 1$ and $s'(x) = c(x) > 0$ on $[0, \beta)$, we get that $s(x) \ge 0$ on $[0, \beta]$. Next we obtain from (1.5) that $s(\beta) = 1$, and by (1.6) we get $s(2\beta) = 2\,s(\beta)\,c(\beta) = 0$. By the choice of $\beta$ and $c(0) = 1$, we are guaranteed that $c(x) > 0$ on $[0, \beta)$. From (1.5) and $s(-x) = -s(x)$ we get that the range of $s(x)$ is $[-1, 1]$. We name $\pi = 2\beta$.

Now $c(2\beta) = c^2(\beta) - s^2(\beta) = 0 - 1 = -1$. Then

$$s(4\beta) = 2s(2\beta)c(2\beta) = 0,$$

and

$$c(4\beta) = c^2(2\beta) - s^2(2\beta) = (-1)^2 - 0 = 1.$$

Also,

$$s(2\beta - x) = s(2\beta)c(-x) + c(2\beta)s(-x) = 0 - s(-x) = s(x).$$

So $s(x) \ge 0$ on $[0, 2\beta]$ and $s(x) > 0$ on $(0, 2\beta)$ (if $s$ had a zero in $(0, 2\beta)$ the above symmetry would imply a zero for $s$ in $(0, \beta)$, and in turn this would imply a zero for $c$ in $(0, \beta)$ by Rolle, a contradiction). Also,

$$s(2\beta + x) = s(2\beta)c(x) + c(2\beta)s(x) = 0 - s(x) = -s(x).$$

So $s(x) \le 0$ on $[2\beta, 4\beta]$. It follows that if $s$ is periodic, its period is at least $4\beta$.

And we have

$$s(x + 4\beta) = s(x)c(4\beta) + c(x)s(4\beta) = s(x) \times 1 + c(x) \times 0 = s(x)$$

and

$$c(x + 4\beta) = c(x)c(4\beta) - s(x)s(4\beta) = c(x) \times 1 - 0 = c(x).$$

So both $s$ and $c$ have period $4\beta = 2\pi$.

The above information allows us to deduce that $(c(t), s(t))$, $0 \le t \le 2\pi$, parametrizes the unit circle. The length of the arc corresponding to the parameter running from $0$ to $x$ is

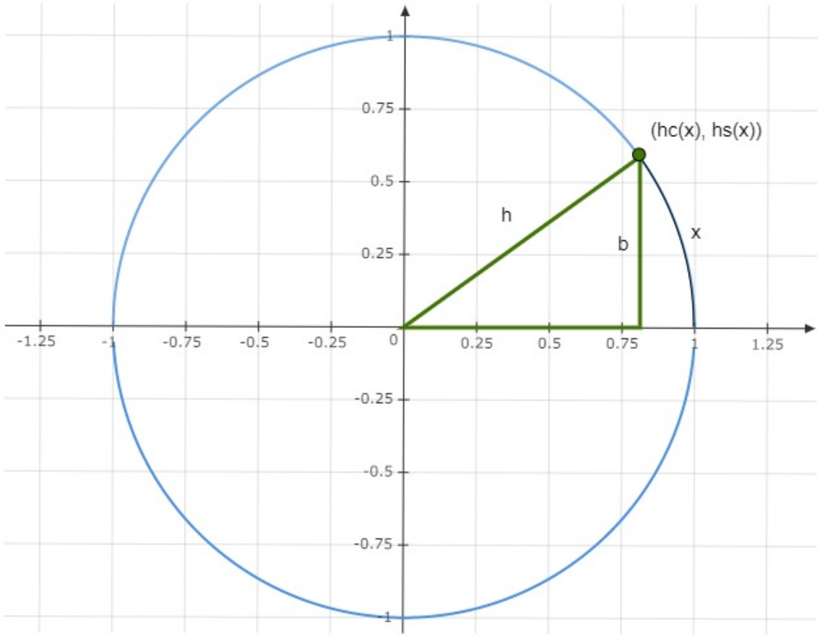
$$\int_0^x \sqrt{c'(t)^2 + s'(t)^2}\,dt = \int_0^x 1\,dt = x.$$

So the angle (in radians) corresponding to the arc joining $(1, 0)$ and $(c(x), s(x))$ is $x$. We also see that $\pi$ as we defined it agrees with the usual $\pi$, for we can parametrize a circle of radius $h$ and centre $(a, b)$ as $(a + h\,c(t), b + h\,s(t))$, $0 \leq t \leq 2\pi'$, where $\pi'$ denotes "our" $\pi$, and we have that the perimeter of the circle is

$$P = \int_0^{2\pi'} \sqrt{h^2 c'(t)^2 + h^2 s'(t)^2}\, dt = \int_0^{2\pi'} h\, dt = 2\pi' h,$$

the usual formula. Hence $\pi' = \pi$.

Now take a square triangle with angle $x$, opposite $b$, and hypotenuse $h$. Fit it in the circle of radius $h$, with the angle $x$ lying at the origin, and the adjacent side on the positive horizontal axis. The picture below is drawn with $h = 1$.




The vertex touching the circle has coordinates $(hc(x), hs(x))$. So $b = hs(x)$, and

$$s(x) = \frac{b}{h} = \sin x.$$

So the "power series sine" is indeed the "triangle sine".

The above process formally defines the sine and cosine, and gives us the basic trigonometric identities

> ***The Basis of all Trigonometry***
>
> *The functions* $\sin x$, $\cos x$ *are everywhere differentiable on* $\mathbb{R}$.

$$\sin^2 x + \cos^2 x = 1, \quad \sin' x = \cos x, \quad \sin(x + 2\pi) = \sin x,$$
$$\sin(x + y) = \sin x \cos y + \cos x \sin y, \qquad \sin(-x) = -\sin x,$$
$$\cos \frac{\pi}{2} = 0, \quad \sin 0 = 0, \quad \cos 0 = 1.$$
$$\frac{\pi}{2} \ \textit{is the smallest positive zero of } \cos x.$$

As hinted by the heading, all trigonometry can be deduced from the above identities (the list is obviously not unique, other choices are possible). Let us work out a few examples of known formulas that do not appear explicitly in the list above. We note first that the zeros of both sine and cosine are isolated. Indeed, the conditions give us that $\cos x > 0$ on $[0, \pi/2)$. As this is the derivative of $\sin x$, we have that $\sin 0 = 0$ and $\sin x$ is increasing on $[0, \pi/2)$. Now

$$\sin\left(x + \frac{\pi}{2}\right) = \sin x \cos \frac{\pi}{2} + \cos x \, \sin \frac{\pi}{2} = \cos x \, \sin \frac{\pi}{2}. \tag{1.8}$$

From $\sin^2 \frac{\pi}{2} + \cos^2 \frac{\pi}{2} = 1$ and $\cos \frac{\pi}{2} = 0$ we know that $\left|\sin \frac{\pi}{2}\right| = 1$. So $\sin x$ is also nonzero on $[\pi/2, \pi)$. We get $\sin \pi = 0$, from

$$\sin \pi = \sin\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = 2 \sin \frac{\pi}{2} \, \cos \frac{\pi}{2} = 0,$$

and then $|\cos \pi| = 1$. Then

$$\sin(x + \pi) = \sin x \cos \pi + \cos x \sin \pi = \sin x \cos \pi.$$

Thus $\sin x \neq 0$ also on $(\pi, 2\pi)$. Combined with $\sin(x + 2\pi) = \sin x$, we have shown that $\sin x = 0$ precisely when $x = k\pi$, $k \in \mathbb{Z}$.

From (1.8) we obtain that the zeros of the cosine are precisely the points $k\pi + \frac{\pi}{2}$, that is

$$\cos x = 0 \iff x = \frac{(2k+1)\pi}{2}, \quad k \in \mathbb{Z}.$$

If we differentiate the identity for the sum with respect to $x$,

$$\cos(x+y) = \frac{d}{dx} \sin(x+y) = \cos x \cos y + (-\sin x) \sin y = \cos x \cos y - \sin x \sin y.$$

Also, using that $\sin(-y) = -\sin y$ and its differentiated version $\cos(-y) = \cos y$,

$$\cos(x - y) = \cos x \cos(-y) - \sin x \sin(-y) = \cos x \cos y + \sin x \sin y.$$

In particular

$$\cos \pi = \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \cos^2 \frac{\pi}{2} - \sin^2 \frac{\pi}{2} = -\sin^2 \frac{\pi}{2} = -1.$$

Therefore

$$\sin(x + \pi) = \sin x \cos \pi = -\sin x.$$

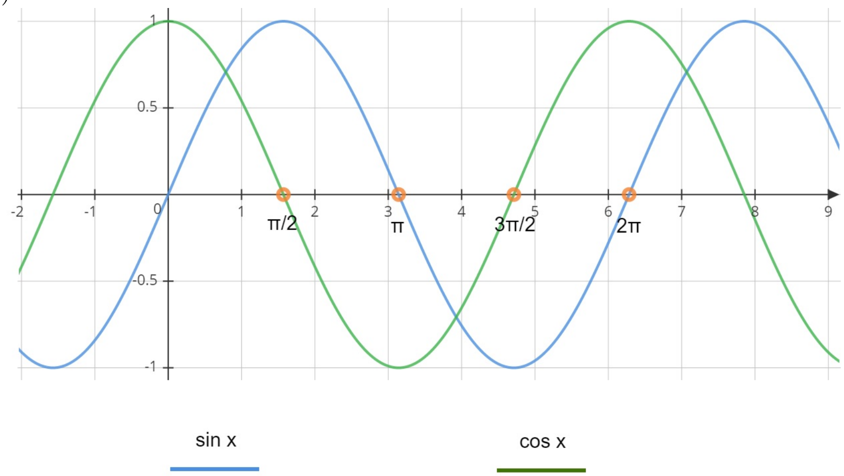
As $\cos 0 = 1$ and $\pi/2$ is the smallest positive zero of $\cos x$, we have that $\cos x > 0$ for $x \in [0, \frac{\pi}{2})$ and so $\cos' \frac{\pi}{2} \leq 0$. That is, $-\sin \frac{\pi}{2} \leq 0$ and then

$$\sin \frac{\pi}{2} = 1.$$

With this last equality, (1.8) becomes

$$\sin\left(x+\frac{\pi}{2}\right)=\sin x\cos\frac{\pi}{2}+\cos x\sin\frac{\pi}{2}=\cos x.$$

The information we have gathered allows us to produce, without a computer nor calculator, the well-known graph (it would just look less pretty if drawn by hand):




Let us keep going a bit more. From

$$1=\sin\frac{\pi}{2}=\sin 2\frac{\pi}{4}=2\sin\frac{\pi}{4}\,\cos\frac{\pi}{4},\qquad 0=\cos\frac{\pi}{2}=\cos 2\frac{\pi}{4}=\cos^2\frac{\pi}{4}-\sin^2\frac{\pi}{4}$$

we easily get

$$\sin\frac{\pi}{4}=\cos\frac{\pi}{4}=\frac{\sqrt{2}}{2}.$$

Using

$$\sin(\pi-x)=\sin\pi\,\cos x-\cos\pi\,\sin x=-(-1)\sin x=\sin x,$$

and

$$\cos(\pi-x)=\cos\pi\,\cos x+\sin\pi\,\sin x=-\cos x,$$

we can calculate

$$\sin\frac{2\pi}{3}=\sin\left(\pi-\frac{\pi}{3}\right)=\sin\frac{\pi}{3},$$

and

$$\cos\left(\frac{2\pi}{3}\right)=\cos\left(\pi-\frac{\pi}{3}\right)=-\cos\frac{\pi}{3}.$$

We get

$$\sin\frac{\pi}{3}=\sin\frac{2\pi}{3}=2\sin\frac{\pi}{3}\,\cos\frac{\pi}{3}.$$

As we know that $\sin\frac{\pi}{3} > 0$, we get

$$\cos\frac{\pi}{3} = \frac{1}{2}, \qquad \sin\frac{\pi}{3} = \sqrt{1-\cos^2\frac{\pi}{3}} = \frac{\sqrt{3}}{2}.$$

Pushing our luck,

$$\frac{\sqrt{3}}{2} = \sin\frac{\pi}{3} = \sin\frac{2\pi}{6} = 2\sin\frac{\pi}{6}\cos\frac{\pi}{6},$$

and

$$\frac{1}{2} = \cos\frac{\pi}{3} = \cos\frac{2\pi}{6} = \cos^2\frac{\pi}{6} - \sin^2\frac{\pi}{6}.$$

A tiny bit of elementary algebra (and solving a biquadratic, where we know that we are looking for the positive roots) gives us

$$\sin\frac{\pi}{6} = \frac{1}{2}, \qquad \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}.$$

These values we have found can be summarized in the following remarkable little table:

| $\mathbf{x}$ | $\sin\mathbf{x}$ | $\cos\mathbf{x}$ |
|---|---|---|
| $\mathbf{0}$ | $\sqrt{0}/2$ | $\sqrt{4}/2$ |
| $\boldsymbol{\pi/6}$ | $\sqrt{1}/2$ | $\sqrt{3}/2$ |
| $\boldsymbol{\pi/4}$ | $\sqrt{2}/2$ | $\sqrt{2}/2$ |
| $\boldsymbol{\pi/3}$ | $\sqrt{3}/2$ | $\sqrt{1}/2$ |
| $\boldsymbol{\pi/2}$ | $\sqrt{4}/2$ | $\sqrt{0}/2$ |

The sine and the cosine of any rational multiple of $\pi$ can be found in principle using these ideas; with the caveat that one can possibly need to find roots of higher degree polynomials and this can lead to equations with no formulas for the roots.

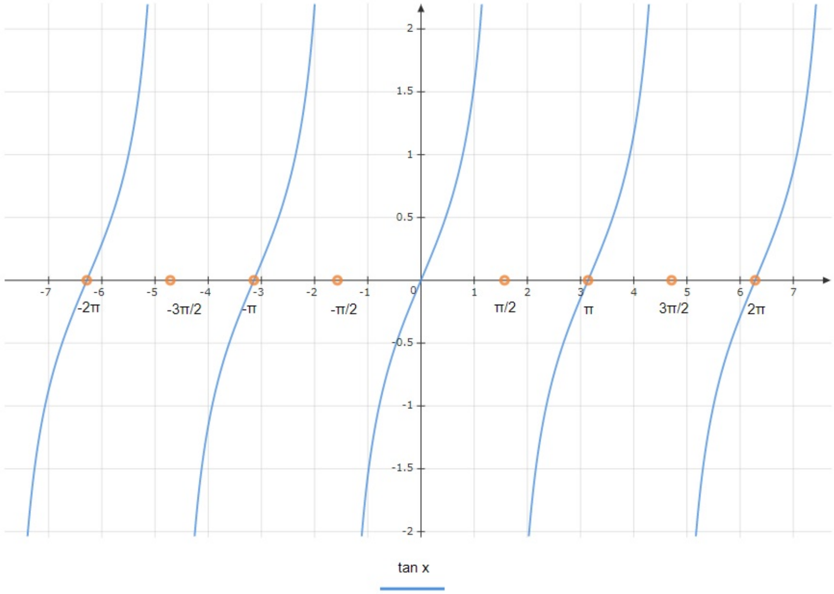
The **tangent function** is $\tan x = \frac{\sin x}{\cos x}$. It fails to exist at those points where $\cos x = 0$, that is $x = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$. We have

$$\tan' x = \frac{\cos^2 x - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

So, where it is defined, the tangent is always increasing. We have $\tan x = 0$ if and only if $\sin x = 0$, so the zeros are $k\pi$, $k \in \mathbb{Z}$. The second derivative is

$$\tan'' x = \frac{2\sin x}{\cos^3 x} = 2\tan x \, \tan' x.$$

Thus the convexity of $\tan x$ agrees with the sign of $\tan x$. With this information it is not hard to come up with the sketch




On the intervals $\left(-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi\right)$, the tangent is increasing and takes all possible real values. So we can consider its inverse function, commonly called $\arctan x$. Concretely, $\arctan x$ is the function defined on all of $\mathbb{R}$ such that $\tan(\arctan x) = x$ for all $x$. As in the inverse function theorem, if $f(g(x)) = x$ we can differentiate to get $1 = f'(g(x))\, g'(x)$. Thus

$$g'(x) = \frac{1}{f'(g(x))}.$$

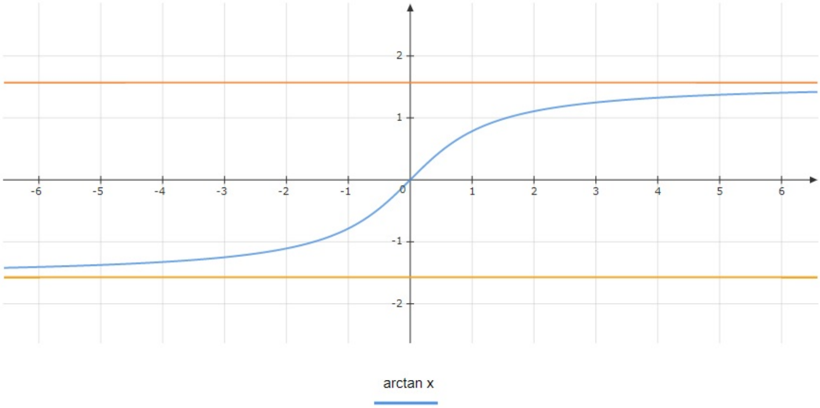
Then

$$\arctan' x = \frac{1}{\frac{1}{\cos^2 \arctan x}} = \cos^2 \arctan x = \frac{1}{1 + \tan^2 \arctan x} = \frac{1}{1+x^2},$$

where we are using the identity

$$\cos^2 x = \frac{1}{1+\tan^2 x},$$

that comes from $\tan^2 x = \sin^2 x / \cos^2 x$ and using $\sin^2 x = 1 - \cos^2 x$. So the inverse tangent function $\arctan x$ is always increasing, goes to $-\frac{\pi}{2}$ at $-\infty$ and to $\frac{\pi}{2}$ at $+\infty$. Its graph looks like




and we also get the well-known integral

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx = \frac{\pi}{2}.$$

**1.4.1. Exercises.**

(1.4.1) Find exact formulas for $\sin\frac{\pi}{5}$ and $\cos\frac{\pi}{5}$.

(1.4.2) Write formulas to express the sine and the cosine in terms of the tangent for $x \in \left[0, \frac{\pi}{2}\right)$. Explain how to adapt the formulas for arbitrary $x$.

(1.4.3) Find an addition formula for the tangent; that is, express $\tan(x+y)$ as a formula on $\tan x$ and $\tan y$.

(1.4.4) For each $x$, let

$$T(x) = \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}.$$

Show that $T(x+y) = T(x)T(y)$.

(1.4.5) Find formulas for $\sin(2x)$ and $\cos(2x)$ in terms of $\tan x$.

(1.4.6) Show that

$$\sin x + \sin y = 2\sin\left(\frac{x+y}{2}\right)\,\cos\left(\frac{x-y}{2}\right),$$

$$\cos x + \cos y = 2\cos\left(\frac{x+y}{2}\right)\,\cos\left(\frac{x-y}{2}\right),$$

and

$$\cos x - \cos y = -2\sin\left(\frac{x+y}{2}\right)\,\sin\left(\frac{x-y}{2}\right).$$

(1.4.7) Show that

$$\tan\frac{x+y}{2} = \frac{\sin x + \sin y}{\cos x + \cos y}$$

whenever the denominator is nonzero.

(1.4.8) Show that
$$\frac{\sin^2 x - \sin^2 y}{\cos^2 x - \cos^2 y} = -1$$
whenever the denominator is nonzero.

(1.4.9) Show that
$$\frac{\sin x + \sin y}{\cos x + \cos y} = -\frac{\cos x - \cos y}{\sin x - \sin y}$$
whenever the denominators are nonzero.

(1.4.10) Show that
$$\{\lambda\, \sin(x+r) :\ \lambda, r \in \mathbb{R}\} = \{\alpha \cos x + \beta \sin x :\ \alpha, \beta \in \mathbb{R}\}.$$

(1.4.11) Use the idea in the proof of Lemma 1.4.1 to show that the initial problem $y' = y$, $y(0) = 1$ has at most one solution.

## 1.5. Complex Numbers

In dealing with polynomial equations centuries ago, mathematicians found themselves needing to use square roots of negative real numbers. This is seen for instance in Cardano's formulas for the roots of a degree three polynomial. In fact, looking at the equation progression in Section 1.3, already the quadratic equation $x^2 + 1 = 0$ has no solution on $\mathbb{R}$. Eventually, centuries ago, mathematicians began to get familiar with using an "imaginary" number $i$ with the property that $i^2 + 1 = 0$. As one wants to use $i$ in the context of adding and multiplying numbers, the set of **complex numbers** arises:
$$\mathbb{C} = \{a + ib :\ a, b \in \mathbb{R}\}$$
is a field that contains $\mathbb{R}$, where the equation $x^2 + 1 = 0$ has a solution (two, in fact, $i$ and $-i$). But, what is $i$? How do we know such a thing exists? It turns out that there are several ways of constructing $\mathbb{C}$, all fairly natural. What we need is a minimum field that contains $\mathbb{R}$ and also an element $i$ with $i^2 = -1$.

At the most basic level we can simply say that $\mathbb{C}$ is $\mathbb{R}^2$, with the operations
$$(a, b) + (c, d) = (a + c, b + d), \qquad (a, b)(c, d) = (ac - bd, ad + bc).$$
Here, we simply see $\mathbb{R}$ as those numbers $(a, 0)$, and the imaginary unit is $i = (0, 1)$.

Another possibility is to define
$$\mathbb{C} = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} :\ a, b \in \mathbb{R} \right\},$$

with the usual matrix addition and multiplication. Here see $\mathbb{R}$ as the real scalar multiples of the identity, and $i = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

A more abstract/algebraic way is to define

$$\mathbb{C} = \mathbb{R}[x]/\langle 1 + x^2 \rangle,$$

the quotient of the ring of real polynomials $\mathbb{R}[x]$ by the ideal generated by the polynomial $1 + x^2$. Here the constant polynomials represent $\mathbb{R}$, and $i$ can be taken to be $x$, since $x^2 + 1 = 0$. Addition and multiplication are given by the respective operations on the representatives.

In all three cases, we are producing a two dimensional field (over $\mathbb{R}$) that contains $\mathbb{R}$ and has an element $i$ with $i^2 = -1$, so all approaches above are equivalent. More concretely, it is straightforward to show that all three approaches yield isomorphic fields, with the isomorphism also preserving the order of $\mathbb{R}$.

The "trick" of using $i$ to get square roots of negative real numbers allows us to solve all quadratic equations. In fact, Cardano and Ferrari's formulas give solutions for any polynomials of degree at most 4 with complex coefficients. But, will we need a bigger set of numbers to have solutions to polynomials equations of higher degree? The answer is no, and turns out to be a spectacular triumph of the introduction of complex numbers:

***1.5.1. Theorem. (Fundamental Theorem of Algebra)***

*Let $p \in \mathbb{C}[x]$ be a polynomial. Then there exists $z_0 \in \mathbb{C}$ with $p(z_0) = 0$. As a consequence, any complex polynomial of degree $n$ can be written as*

$$p(z) = a\,(z - z_1) \cdots (z - z_n), \qquad a, z_1, \dots, z_n \in \mathbb{C}.$$

There are no known elementary proofs of the Fundamental Theorem of Algebra, particularly in the sense that there is no purely algebraic proof. All proofs seem to use analysis one way or another. We will provide one such proof in Corollary 3.7.5.

Given $z = a + ib \in \mathbb{C}$, its **conjugate** is the complex number $\bar{z} = a - ib$. The **absolute value** or **modulus** is the non-negative number $|z| = (\bar{z}z)^{1/2} = (a^2 + b^2)^{1/2}$. The **real part** of $z$ is the real number $\operatorname{Re} z = a = \frac{z+\bar{z}}{2}$. The **imaginary part** of $z$ is the real number $\operatorname{Im} z = b = \frac{z-\bar{z}}{2i}$.

As with the real line, the absolute value gives us a notion of distance in $\mathbb{C}$, that allows us to do topology. Namely, we say that $z_n \to z$ if $|z_n - z| \to 0$. As the latter limit occurs in $\mathbb{R}$, we are defining the topology in $\mathbb{C}$ by means of the usual one in $\mathbb{R}$; in particular, convergence of sequences of real numbers in $\mathbb{C}$ is precisely convergence in $\mathbb{R}$.

Since we have a field and we can take limits, series make sense. For any $z \in \mathbb{C}$, consider the sequence of partial sums

$$s_n = \sum_{k=0}^{n} \frac{z^k}{k!}.$$

Since, for $n > m$,

$$|s_n - s_m| = \left| \sum_{k=m+1}^{n} \frac{z^k}{k!} \right| \leq \sum_{k=m+1}^{n} \frac{|z|^k}{k!}.$$

The convergence of the usual exponential series in $\mathbb{R}$ then guarantees that the sequence $\{s_n\}$ is Cauchy. So its limit exists and we can write

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

With the same proofs as in the real case we can show the basic properties

$$e^0 = 1, \qquad e^{z+w} = e^z e^w, \quad z, w \in \mathbb{C}.$$

In particular, $e^{-z} = 1/e^z$ and so $e^z \neq 0$ for all $z$. By the usual properties of power series, the function $f(z) = e^z$ is infinitely differentiable; namely, both the series for $e^z$ and its derivative (which are the same) converge uniformly on the interior of any disk, and we can differentiate term by term. So, what changes when using a complex argument for the exponential, as opposed to real? Suppose that $z = i\theta$, with $\theta \in \mathbb{R}$. Then, using that $i^2 = -1$, $i^3 = -i$, and $i^4 = 1$,

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!} = \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} = \cos\theta + i \sin\theta.$$

This allows us to write, if $z = a + ib$ with $a, b \in \mathbb{R}$,

$$e^{a+ib} = e^a e^{ib} = e^a(\cos b + i \sin b).$$

We obtain in particular the notable formula that relates the most famous numbers:

$$e^{i\pi} + 1 = 0.$$

More importantly, it gives us the **polar form** of a complex number; we can always write $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$, with $r \geq 0$ and $\theta \in \mathbb{R}$. We also see that the exponential function is not injective on $\mathbb{C}$. The polar form works very nicely with multiplication, for

$$(re^{i\theta})(se^{i\gamma}) = (rs)e^{i(\theta+\gamma)}.$$

An application of this is to find the $n^{\text{th}}$ **roots of unity**. Namely, the $n$ solutions of the equation $\omega^n = 1$ are

$$\omega_k = e^{2i\pi k/n}, \qquad\qquad k = 0, \ldots, n-1.$$

Other relations that follow directly from the definitions of the sine, cosine, and exponential are

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

It is worth mentioning a common pitfall with complex numbers' arithmetic. As constructed, the addition and multiplication in $\mathbb{C}$ satisfy all the properties one knows from the reals (commutativity, distributivity, etc.) But that does not mean that any formula that works for real numbers will work for complex numbers. One such rule one has to be careful with, is

$$(a^b)^c = a^{bc}.$$

This holds whenever $a, b, c$ are real numbers with $a > 0$. The restriction $a > 0$ is necessary for $a^b$ to make sense (usually defined as $a^b = e^{b \log a}$) for arbitrary $b$. Now, even before moving towards $\mathbb{C}$, this rule has problems; for instance

$$((-1)^2)^{1/2} = 1^{1/2} = 1 \neq -1 = (-1)^{2\times 1/2}.$$

So care is required when dealing with complex numbers algebraically. The basic idea for solving equations is that "doing the same thing to both sides of the equality will preserve the equality". But in $\mathbb{C}$ we have $e^0 = e^{2\pi i}$. One would then expect $(e^0)^{1/2} = (e^{2\pi i})^{1/2}$, but the reality is that the square root is not obviously defined everywhere (not even in $\mathbb{R}$, as shown in the example above); we have $e^{0\times 1/2} = e^0 = 1$, while $e^{2\pi i/2} = e^{\pi i} = -1$.

**1.5.1. Exercises.**

(1.5.1) Without using series, show that a nonzero complex number $z$ can be written in a unique way as $z = r(\cos\theta + i\sin\theta)$ with $r > 0$ and $\theta \in [0, 2\pi)$.

(1.5.2) Still without using series, follow up from Exercise 1.5.1 by showing that if we formally use the notation $e^{i\theta}$ for the complex number $\cos\theta + i\sin\theta$, then $e^{i(\theta_1+\theta_2)} = e^{i\theta_1}e^{i\theta_2}$.

(1.5.3) Let $z \in \mathbb{C}$ be nonzero. Show that if $z = r(\cos\theta + i\sin\theta)$, then

$$z^n = r^n(\cos n\theta + i\sin n\theta).$$

(1.5.4) Show that the equation $z^n = 1$ has precisely $n$ distinct solutions, that can be expressed as

$$\omega_k = e^{2i\pi k/n}, \qquad k = 0, \ldots, n-1.$$

(1.5.5) Solve the equation $(z+1)^5 = z^5$.

(1.5.6) Show that $e^{z+w} = e^z e^w$ by using the series expansion.

(1.5.7) Show that $e^{z+w} = e^z e^w$ by showing that $e^{z+w}$ is the unique solution of the initial value problem $y'(z) = y(z)$, $y(0) = e^w$.

## 1.6. Cardinality

A far reaching idea, due to Cantor, is to notice that intuitively two sets have the same amount of elements if and only if there is a bijection between them. We will write $|A| = |B|$ to denote the existence of a bijection $A \to B$. The relation $A \sim B$ if and only if $|A| = |B|$ is an equivalence relation.

**1.6.1. Definition.**

*We say that two sets $A, B$ have **equal cardinality** (or that they are **equipotent**) if $|A| = |B|$, that is if there exists a bijection $A \to B$.*

*We write $|A| \leq |B|$ if there exists an injection $A \to B$.*

*A set $A$ is **infinite** if there exists $B \subsetneq A$ with $|A| = |B|$.*

*A set $A$ is **finite** if it is not infinite.*

We can immediately check that $\mathbb{N}$ is infinite, since $|2\mathbb{N}| = |\mathbb{N}|$ via the bijection $\gamma : \mathbb{N} \to 2\mathbb{N}$, $\gamma(n) = 2n$. And from this we can show that if $A \supset \mathbb{N}$, then $A$ is infinite. Indeed, we have $A = (A \setminus \mathbb{N}) \cup \mathbb{N}$. Then we have $|(A \setminus \mathbb{N}) \cup 2\mathbb{N}| = |A|$ via the bijection $\gamma : A \to (A \setminus \mathbb{N}) \cup 2\mathbb{N}$,

$$\gamma(a) = \begin{cases} a, & a \in A \setminus \mathbb{N} \\ 2a, & a \in \mathbb{N} \end{cases}$$

The same idea works to check that any set that contains an infinite set is infinite. Surprisingly, it requires a bit of work to characterize finite sets.

***1.6.2. Proposition.***

*A non-empty set $F$ is finite if and only if there exists $m \in \mathbb{N}$ such that $|F| = |\{1, \ldots, m\}|$.*

*Proof.* Suppose first that $F$ is finite. Let

$$\mathcal{C} = \{(g, G) : \ g : F \to \mathbb{N}, \ G \subset F, \ g(G) = \operatorname{ran} g, \ g|_G \ \text{ injective}\},$$

with the order $(g_1, G_1) \leq (g_2, G_2)$ if $G_1 \subset G_2$ and $g_2|_{G_1} = g_1$. Note that $\mathcal{C} \neq \varnothing$ because we can take a fixed $z \in F$, $G = \{z\}$, and $g(z) = 1$. If $\{(g_j, G_j)\}$ is a chain, then we construct $G = \bigcup_j G_j$ and define $g : F \to \mathbb{N}$ by $g(x) = g_j(x)$ if $x \in G_j$ (this is well-defined by the compatibility given by the chain relation) and $g = 0$ on $F \setminus G$. If $x, y \in G$ with $g(x) = g(y)$, there exists $j$ such that $x, y \in G_j$ and so $g_j(x) = g(x) = g(y) = g_j(y)$ and $x = y$

by the injectivity of $g_j$, so $g$ is injective. By Zorn's Lemma there exists a maximal element $(h, H) \in \mathcal{C}$. This function $h$ cannot be surjective, because then $|H| = |\mathbb{N}|$, and we can deduce that $F$ is infinite, a contradiction (namely, map $H$ to a proper subset and $F \setminus H$ to itself to obtain a bijection between $F$ and a proper subset). So $h$ is not surjective. But then if $F \setminus H \neq \varnothing$, take $z \in F \setminus H$, and $m \in \mathbb{N} \setminus h(H)$, and we can extend $h$ to $H \cup \{z\}$ by $h(z) = m$ and it is still injective. The maximality of $(h, H)$ then implies that $H = F$. So we have an injective function $h : F \to \mathbb{N}$, which is not surjective. If $h(F)$ is not bounded then it is infinite, and we can repeat the argument to imply that $F$ is infinite, a contradiction; thus, $h(F)$ is a bounded subset of $\mathbb{N}$. That is, $h(F) = \{n_1, \ldots, n_m\} \subset \mathbb{N}$ for some $m \in \mathbb{N}$. The composition of $h$ with the inverse of $n_*$ gives a bijection between $F$ and $\{1, \ldots, m\}$.

The converse is the easy part. We want to show that $R_n = \{1, \ldots, m\}$ is finite. This, after considering reorderings, amounts to show that there is no bijection between $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$ if $n \neq m$ (Exercise 1.6.2). □

Let us now focus on examples of cardinality.

*1.6.3. Example.* We saw above that $|2\mathbb{N}| = |\mathbb{N}|$. And there was nothing particular about 2 there, we could have used any other integer. What about $\mathbb{N} \cup \{0\}$? We can define $\gamma : \mathbb{N} \cup \{0\} \to \mathbb{N}$ by $\gamma(n) = n + 1$; this is a bijection so $|\mathbb{N} \cup \{0\}| = |\mathbb{N}|$.

*1.6.4. Example.* An interesting fact about the equality $|2\mathbb{N}| = |\mathbb{N}|$ is that the remaining part of $\mathbb{N}$, the odd integers, is also equipotent with $\mathbb{N}$. Indeed, we can define $\gamma : \mathbb{N} \to 2\mathbb{N} + 1$ by $\gamma(n) = 2n - 1$, and this is a bijection. So $\mathbb{N}$ can be partitioned into two "halves" each the same size as the full set. We can use this to deal with another well-known infinite set, $\mathbb{Z}$, the integers. Let $\gamma : \mathbb{Z} \to \mathbb{N}$,

$$\gamma(n) = \begin{cases} 2n + 2, & n \geq 0 \\ -2n - 1, & n < 0 \end{cases}$$

The reader should check that this is a bijection, so $|\mathbb{Z}| = |\mathbb{N}|$.

*1.6.5. Example.* What about the rationals $\mathbb{Q}$? We will now show that $|\mathbb{Q}^+| = |\mathbb{Q}|$. We can use the following idea. We partition $\mathbb{Q}$ into intervals of length one between integers, and then we implement a version of the bijection between $\mathbb{Z}$ and $\mathbb{N}$. Concretely, for $n \geq 0$ let us map $[n, n+1)$ to $[2n, 2n+1)$ and $[-n-1, -n)$

to $[2n+1, 2n+2)$. With $n(q) = |\lfloor q \rfloor|$, we put

$$\gamma(q) = \begin{cases} n(q) + q, & q \geq 0 \\ n(q) + 1 - q, & q < 0 \end{cases}$$

The reader is encouraged to calculate the inverse explicitly (Exercise 1.6.4).

*1.6.6. Example.* There are infinitely many rationals between any two integers, so it certainly feels like that there are a lot more rationals than integers. But it is not hard to order the rationals in such a way that they form a sequence. That is, $|\mathbb{Q}| = |\mathbb{N}|$. Here is one way to do it. First we will show that $|\mathbb{Q}^+| = |\mathbb{N} \cup \{0\}|$, and then we will be done by using Example 1.6.5. Since there are infinitely many prime numbers, we can enumerate them as $\{p_n\}$. They key observation is that any positive rational is of the form $p_{j_1}^{r_1} \cdots p_{j_k}^{r_k}$ for primes $p_{j_1}, \dots, p_{j_k}$ and $r_1, \dots, r_k \in \mathbb{Z}$, and that this decomposition is unique. This follows directly from the Fundamental Theorem of Arithmetic, i.e. that any positive integer is of the form $p_{j_1}^{r_1} \cdots p_{j_k}^{r_k}$ with $r_j \geq 0$ for all $j$. If we now fix a bijection $\rho_0 : \mathbb{N} \cup \{0\} \to \mathbb{Z}$, we can define

$$\gamma(p_{j_1}^{r_1} \cdots p_{j_k}^{r_k}) = p_{j_1}^{\rho(r_1)} \cdots p_{j_k}^{\rho(r_k)}.$$

Then $|\mathbb{N}| = |\mathbb{Q}^+| = |\mathbb{Q}|$. There are many other ways to do this, with different ways of laying the rationals as a list. We will not worry too much about it, because we will have yet another argument in Example 1.6.15.

*1.6.7. Example.* Let $(a, b) \subset \mathbb{R}$ be an interval. Then $\gamma : (a, b) \to (0, 1)$ given by $\gamma(x) = (x - a)/(b - a)$, is a bijection. Given another interval $(c, d)$, we could of course compose bijections to get an explicit function, but we can also simply use that equipotency is an equivalence relation (which amounts to the fact that a composition of bijections is a bijection) to say $|(c, d)| = |(0, 1)| = |(a, b)|$.

*1.6.8. Example.* Could we compare $\mathbb{R}$ with an interval? We obviously have $|(0, 1)| \leq |\mathbb{R}|$ since $(0, 1) \subset \mathbb{R}$ so there is a clear injection. But consider $\arctan : \mathbb{R} \to \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$. This is a bijection, so $|\mathbb{R}| = \big|\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)\big| = |(0, 1)|$. We can also use the exponential to show that $|\mathbb{R}| = |(0, \infty)|$. And instead of arctan and changing intervals we can also use the map $g : \mathbb{R} \to (0, 1)$ given by $g(t) = \frac{1}{2}\big(\frac{t}{|t|+1} + 1\big)$.

*1.6.9. Example*. What happens if we remove a point from an infinite set? We will see a general way to do it in Corollary 1.6.18, but let us show here that $|\mathbb{R} \setminus \{0\}| = |\mathbb{R}|$. We somehow need to "make room" to put 0 somewhere else. The idea is to use the natural numbers for that. That is, we can take $\gamma : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by

$$\gamma(r) = \begin{cases} r+1, & r \in \mathbb{N} \cup \{0\} \\ r, & \text{otherwise} \end{cases}$$

We "shift over the naturals" and leave the rest alone. Thus a hole is created at zero. Since exchanging 0 with a fixed $r \in \mathbb{R}$ and leaving the rest alone is a bijection, we immediately have $|\mathbb{R} \setminus \{r\}| = |\mathbb{R}|$. We cannot expect any of these bijections to be continuous, as continuous functions map connected sets to connected sets.

*1.6.10. Example*. Given $a < b < c < d \in \mathbb{R}$, we can now show that $|\mathbb{R}| = |(a, b) \cup (c, d)|$. If we were given this before the previous examples, it would have probably looked hard to find a bijection, or even to determine if one exists. But now we can consider $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$, and define explicitly $\gamma : \mathbb{R} \setminus \{0\} \to (a, b) \cup (c, d)$ by

$$\gamma(t) = \begin{cases} \dfrac{a-b}{1-t} + b, & t < 0 \\ \dfrac{c-d}{1+t} + d, & t > 0 \end{cases}$$

where the idea is to map the negative reals to $(a, b)$ and the positive reals to $(c, d)$ (and, of course, there are many ways to do this). Now we have $|(a, b) \cup (c, d)| = |\mathbb{R} \setminus \{0\}| = |\mathbb{R}|$. And combining examples we now know $|(0, 1)| = |(0, 1) \cup (1, 2)|$ and we can even work out the bijection explicitly (Exercise 1.6.3).

*1.6.11. Example*. We can use the same idea we used to remove a point from $\mathbb{R}$ to remove a point from other sets. For instance consider $[0, 1)$ and $(0, 1)$. As long as we have a sequence inside $[0, 1)$, we can use it to "hide" 0. Concretely, we can define a bijection $\gamma : [0, 1) \to (0, 1)$ by

$$\gamma(t) = \begin{cases} \frac{1}{2}, & t = 0 \\ \frac{1}{n+1}, & t = \frac{1}{n},\ n \in \mathbb{N} \setminus \{1\} \\ t, & \text{otherwise} \end{cases}$$

Then $|[0,1)| = |(0,1)|$. We can push the idea, of course. Here is a bijection $\gamma : [0,1] \to (0,1)$.

$$\gamma(t) = \begin{cases} \frac{1}{2}, & t = 0 \\ \frac{1}{3}, & t = 1 \\ \frac{1}{n+2}, & t = \frac{1}{n},\ n \in \mathbb{N} \setminus \{1\} \\ t, & \text{otherwise} \end{cases}$$

*1.6.12. Example.* The key idea in the last few examples is that if we have a sequence inside our set, we can use it to "shift an element out". Could we do it with infinitely many elements? Let $I = [0,1] \setminus \mathbb{Q}$, the irrational numbers in $[0,1]$. Let us show that $|I| = |[0,1]|$. Let $\{q_n\} \subset [0,1]$ be an enumeration of $\mathbb{Q} \cap [0,1]$, and let $\{p_n\}$ be an enumeration of the prime numbers. We have that $\sqrt{p_n}$ is irrational for each $n$, so $1/\sqrt{p_n} \in I$ for each $n$. Then we can define $\gamma : [0,1] \to I$ by

$$\gamma(t) = \begin{cases} \dfrac{1}{\sqrt{p_{2n}}}, & t = q_n \\ \dfrac{1}{\sqrt{p_{2n-1}}}, & t = \frac{1}{\sqrt{p_n}} \\ t, & \text{otherwise} \end{cases}$$

What we are doing here is using the sequence $\{a_n\}$ with $a_n = 1/\sqrt{p_n}$ to "hide" the rationals. We use the even indices to insert the rationals, and the odd indices to re-accommodate the $a_n$. We have shown that $|I| = |[0,1]|$.

Cardinality works fine in the following intuitive sense: if $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$. Easier said than done! The Schröder–Bernstein theorem is very useful to show that sets are equipotent since it is often easier to construct injections than bijections.

***1.6.13. Theorem. (Schröder–Bernstein)***

*Let $A, B$ be sets such that $|A| \leq |B|$ and $|B| \leq |A|$. Then $|A| = |B|$.*

*Proof.* Let $f : A \to B$, $g : B \to A$ be injective. Define subsets of $A$ as follows:

$$A_0 = A, \qquad B_1 = g(B), \qquad A_{k+1} = g \circ f(B_{k+1}), \qquad B_{k+1} = g \circ f(A_k).$$

We have $B_1 \subset A_0$, and $A_1 = g(f(B_1)) \subset g(f(A_0)) \subset g(B) = B_1$. Now assume for induction that $B_k \subset A_{k-1}$ and $A_k \subset B_k$. Then

$$B_{k+1} = g(f(A_k)) \subset g(f(B_k)) = A_k.$$

And

$$A_{k+1} = g(f(B_{k+1})) \subset g(f(A_k)) = B_{k+1}.$$

Thus

$$A_0 \supset B_1 \supset A_1 \supset B_2 \supset A_2 \supset \cdots$$

Let $h = g \circ f \circ g \circ f$. We claim that $h$ is a bijection between $A_k \setminus B_{k+1}$ and $A_{k+1} \setminus B_{k+2}$. Indeed, if $x \in A_k \setminus B_{k+1}$, then $h(x) \in A_{k+1}$. If $h(x) \in B_{k+2}$, then $h(x) = g(f(a))$ for some $a \in A_{k+1}$. As $g \circ f$ is injective, $a = g(f(x)) \in g \circ f(A_k) = B_{k+1}$, a contradiction. So $h$ maps $A_k \setminus B_{k+1}$ into $A_{k+1} \setminus B_{k+2}$. And $h$ is onto: given any $a \in A_{k+1} \setminus B_{k+2}$, since $A_{k+1} = h(A_k)$ there exists $a_0 \in A_k$ with $a = h(a_0)$. If $a_0 \in B_{k+1}$, then $a = h(a_0) \in B_{k+2}$, a contradiction. So $a_0 \in A_k \setminus B_{k+1}$, and $h$ is onto $A_{k+1} \setminus B_{k+2}$, giving us

$$|A_k \setminus B_{k+1}| = |A_{k+1} \setminus B_{k+2}|.$$

Let

$$C = \bigcap_k A_k = \bigcap_k B_k.$$

Now we can write the disjoint unions

$$A = C \cup \bigcup_{k=0}^{\infty} \big(A_k \setminus B_{k+1}\big) \cup \bigcup_{k=1}^{\infty} \big(B_k \setminus A_k\big),$$

$$B_1 = C \cup \bigcup_{k=1}^{\infty} \big(A_k \setminus B_{k+1}\big) \cup \bigcup_{k=1}^{\infty} \big(B_k \setminus A_k\big).$$

Then $\gamma : A \to B_1$ given by

$$\gamma(x) = \begin{cases} x, & x \in C \cup \bigcup_{k=1}^{\infty} B_k \setminus A_k, \\ h(x), & x \in \bigcup_{k=0}^{\infty} A_k \setminus B_{k+1} \end{cases}$$

is a bijection. Thus $|B| = |B_1| = |A|$. $\square$

*1.6.14. Example*. We already know that $|(0,1)| = |[0,1)|$. The bijection required a neat idea to be constructed, namely the shifting over a countable subset. Without that idea, we could have used Schröder–Bernstein. We have $(0,1) \subset [0,1)$ as an obvious injection. But we can also map $[0,1)$ injectively to $(0,1/2] \subset (0,1)$ by using $\gamma(t) = (1-t)/2$. Thus $|(0,1)| = |[0,1)|$ by Schröder–Bernstein.

*1.6.15. Example*. Here is a less trivial use of Schröder–Bernstein. Consider the Cartesian product $\mathbb{N} \times \mathbb{N}$. Instead of stopping to think of a bijection with $\mathbb{N}$, we can notice that we can map $\mathbb{N} \to \mathbb{N} \times \{1\}$, so $|\mathbb{N}| \leq |\mathbb{N} \times \mathbb{N}|$ as expected. But we can map $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ injectively by $(n,m) \longmapsto 2^n 3^m$ (of course, other primes can be used if desired). Thus $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$ by Schröder–Bernstein. With the same idea, $|\mathbb{Z}\times\mathbb{Z}| = |\mathbb{Z}|$, or we can do first $|\mathbb{Z}\times\mathbb{Z}| = |\mathbb{N}\times\mathbb{N}|$ by taking a bijection $\eta : \mathbb{Z} \to \mathbb{N}$ and mapping $(n, m) \longmapsto (\eta(n), \eta(m))$. If we consider the rationals $\mathbb{Q}$, they are made out of pairs of integers, but many pairs produce the same rational. It is not terribly hard to find a bijection as we did in Example 1.6.6, but we can simply notice that we can map $\mathbb{Q} \to \mathbb{Z}\times\mathbb{Z}$ injectively by

$$\gamma(a/b) = (a, b),$$

where for the function to be well-defined we assume that $a/b$ is reduced and $b > 0$. Then $|\mathbb{N}| \leq |\mathbb{Q}| \leq |\mathbb{Z}\times\mathbb{Z}| = |\mathbb{Z}| = |\mathbb{N}|$, and so $|\mathbb{Q}| = |\mathbb{N}|$ by Schröder–Bernstein.

*1.6.16. Remark* . While we cannot emphasize enough the importance of Schröeder–Benrstein's Theorem, it is not really needed to show that $|\mathbb{N}\times\mathbb{N}| = |\mathbb{N}|$. Indeed, it is not hard to come up with an explicit bijection. Concretely, $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ given by

$$g(m, n) = 2^{m-1}(2n - 1)$$

is a bijection. The bijections with $\mathbb{Q}$ are also doable explicitly, but they are more convoluted to express.

*1.6.17. Example* ("A countable union of countable sets is countable"). Let $\{A_n\}_{n\in\mathbb{N}}$ be sets with $|A_n| = |\mathbb{N}|$ for each $n$. Then

$$\Big|\bigcup_n A_n\Big| = |\mathbb{N}|.$$

Indeed, we have bijections $\gamma_n : A_n \to \mathbb{N}$. Let $\{p_n\}$ be an enumeration of the prime numbers. The sets $P_n = \{p_n^k :\ k \in \mathbb{N}\} \subset \mathbb{N}$ are disjoint, and $|P_n| = |\mathbb{N}|$ for each $n$. By mapping $\mathbb{N}$ to $A_1$ we get $|\mathbb{N}| \leq |\bigcup_n A_n|$. And given $a \in \bigcup_n A_n$, let $m = \min\{n :\ a \in A_n\}$ and map $a \longmapsto p_m^{\gamma_m(a)}$. This defines an injection, so $|\bigcup_n A_n| \leq |\mathbb{N}|$ and thus

$$\Big|\bigcup_n A_n\Big| = |\mathbb{N}|.$$

The exact same idea can be used to show that a countable union of finite sets is countable, when one notices that only the injectivity of the $\gamma_m$ is relevant the argument.

Another consequence of Schröder–Bernstein is the following.

**1.6.18. Corollary.**

*Let $A$ be an infinite set, and $F \subset A$ be finite. Then $|A| = |A \setminus F|$.*

*Proof.* Assume first that $|F| = \{a_0\}$. Since $A$ is infinite, there exists $B \subsetneq A$ with $|B| = |A|$. Let $b_0 \in A \setminus B$. Then

$$|A| = |B| \leq |B \cup [(A \setminus B) \setminus \{b_0\}]| = |A \setminus \{b_0\}| \leq |A|.$$

Then Schröder–Bernstein gives us $|A \setminus \{b_0\}| = |A|$. Next define a bijection $\gamma : A \setminus \{b_0\} \to A \setminus \{a_0\}$ by

$$\gamma(a) = \begin{cases} a, & a \neq a_0 \\ b_0, & a = a_0 \end{cases}$$

Then $\gamma$ is a bijection and we have $|A| = |A \setminus \{b_0\}| = |A \setminus \{a_0\}|$. For arbitrary finite $F$, we proceed by induction. Note that $A \setminus \{b_0\}$ is still infinite because it is equipotent with $B$, which is infinite (as it is equipotent with $A$). □

It would have been tempting to use the "countable shift" idea that we have used previously. But the problem is that we still don't know that $|\mathbb{N}| \leq |A|$ for any infinite set $A$. Until now.

**1.6.19. Corollary.**

*Let $A$ be an infinite set. Then $|A| \geq |\mathbb{N}|$.*

*Proof.* By Corollary 1.6.18, there exists $a_1 \in A$ such that $|A| = |A \setminus \{a_1\}|$. Applying Corollary 1.6.18 again, there exists $a_2 \in A \setminus \{a_1\}$ with $|A \setminus \{a_1, a_2\}| = |A \setminus \{a_1\}|$. Inductively we get a set of distinct elements $A_0 = \{a_1, a_2, \ldots\} \subset A$. Then $|\mathbb{N}| = |A_0| \leq |A|$. □

*1.6.20. Remark.* Since we are not being particularly careful with the use of the axioms of set theory—though we still use arguments that are reasonable to mathematicians and suitable to formalize in the common ZFC framework—axioms will not be commonly mentioned. The proper wording in the proof of Corollary 1.6.19 would be that we construct $A_0$ using recursion.

With the natural intuition that "infinite is infinite" one would expect to be able to show the reverse inequality in Corollary 1.6.19, but we are in for a surprise. What we have so far, from Corollary 1.6.19 is that $\mathbb{N}$ is the smallest infinite set, and every infinite set contains a copy of $\mathbb{N}$.

**1.6.21. Definition.**

*We say that a set $A$ is* ***countable*** *if there exists an injective function $\alpha : A \to \mathbb{N}$.*

The idea is that $A$ has at most as many elements as $\mathbb{N}$. Under this definition, every finite set is countable, and every subset of $\mathbb{N}$, finite or not, is countable. So far nothing about infinities is too outrageous, since it should not be much of a surprise that a subset of an infinite set can be put into bijection with the bigger set. But it turns out that is not always the case. In the long list of examples we covered, we showed that lots of sets are equipotent with $\mathbb{N}$, and lots of sets are equipotent with $\mathbb{R}$; but we never linked these two sets. That was no accident!

***1.6.22. Theorem. (Cantor's Diagonal Argument)***

*The set $\mathbb{R}$ of real numbers is not countable. That is, $|\mathbb{N}| < |\mathbb{R}|$.*

*Proof.* Suppose that $f : \mathbb{R} \to \mathbb{N}$ is injective. Define $g : \mathbb{N} \to \mathbb{R}$ by $g(f(r)) = r$ and $g(n) = 0$ if $n \notin f(\mathbb{R})$. Then $g$ is surjective. For any $r \in \mathbb{R}$ and $k \in \mathbb{N}$, let $d(k, r)$ be the $k^{\text{th}}$ digit to the right of the period of $r$; for example, $d(2, 5.278) = 7$. To have $d$ well-defined, we need to fix a single representation for each real number, to avoid things like $0.2 = 0.1999\cdots$. We assume the second representation, that is any rational non-periodic number is written ending with an infinite sequence of nines. Let $h : \{0, \ldots, 9\} \to \{0, \ldots, 9\}$ such that $h(b) \neq b$ for all $b = 0, \ldots, 9$; for instance $h(b) = 5$ if $b \neq 5$ and $h(5) = 4$. Then consider the number

$$s = \sum_{k=1}^{\infty} \frac{h(d(k, g(k)))}{10^k} \in [0, 1].$$

As $s$ does not have a single 9 in its decimal representation, there is no possible ambiguity. By construction, the first digit of $s$ differs from the first digit of $g(1)$; the second digit of $s$ differs from the second digit of $g(2)$; and so on. Thus $s \notin g(\mathbb{N})$, contradicting the fact that $g$ was surjective. The conclusion is that $f$ cannot exist, and $\mathbb{R}$ is not countable. $\square$

*1.6.23. Remark.* A more intuitive way of laying out Cantor's diagonal argument is by assuming that the real numbers in $(0, 1)$ form a sequence $\{x_n\}$, and writing

$$\begin{aligned} x_1 &= 0.x_{11}x_{12}x_{13}\cdots \\ x_2 &= 0.x_{21}x_{22}x_{23}\cdots \\ x_3 &= 0.x_{31}x_{32}x_{33}\cdots \\ &\vdots \end{aligned}$$

where each $x_{kj}$ is a digit. Then one forms

$$z = 0.z_1 z_2 z_3 \cdots$$

where $z_1 \neq x_{11}$, $z_2 \neq x_{22}$, $z_3 \neq x_{33}$, etc. Then $z$ is distinct from all $x_k$, leading to a contradiction. This argument is exactly what was written more formally in the proof of Theorem 1.6.22. We prefer the more formal form of the argument because, simple as it is, the diagonal argument is controversial for some people; and there is the possibility that their objections are based on the informality with which the argument is usually presented.

*1.6.24. Remark* . For another proof of the uncountability of $\mathbb{R}$, based on topological ideas, see Corollary 1.8.22. A third proof, using Lebesgue's outer measure, can be found on Remark 2.3.13.

So not all sets are countable. Theorem 1.6.22 implies that $\mathbb{R}$ is somehow "more infinite" than $\mathbb{N}$. Cantor's idea proved to be far reaching; it lies behind such amazing results as Gödel's Incompleteness Theorems and Church/Turing's solution of the Halting Problem.

The nontrivial part on the next result is the fact that any infinite subset of $\mathbb{N}$ is equipotent with $\mathbb{N}$.

***1.6.25. Proposition.***

*An infinite set $A$ is countable if and only if $|A| = |\mathbb{N}|$.*

*Proof.* If $|A| = |\mathbb{N}|$, then $A$ is countable by definition. Conversely, if there exists an injection $A \to \mathbb{N}$, to show that $|A| = |\mathbb{N}|$ amounts to showing that $|B| = |\mathbb{N}|$ for any infinite subset $B$ of $\mathbb{N}$ (where $B$ is the range of an injection $A \to \mathbb{N}$). For this, because $\mathbb{N}$ is well-ordered, $B$ has a first element, that we denote by $g(1)$. Then $B \setminus \{g(1)\}$ has a first element, that we denote by $g(2)$. Inductively, given $g(1), \dots, g(k)$, we define $g(k+1)$ as the first element of $B \setminus \{g(1), \dots, g(k)\}$. The function $g : \mathbb{N} \to B$ is injective by definition. As $B$ is infinite, the sequence $g(1) < g(2) < \cdots$ is unbounded. If $b \in B \setminus g(\mathbb{N})$, there would exist $k$ such that $g(k) > b$; but this implies that $b = g(j)$ for some $j < k$, as when $g(k)$ was chosen it was the least element in $B \setminus \{g(1), \dots, g(k-1)\}$. This contradiction shows that $g$ is surjective. Thus $|B| = |\mathbb{N}|$, and so $|A| = |B| = |\mathbb{N}|$. $\square$

Here is another use of the "diagonal trick". It implies the amazing fact that there are infinitely many different infinities.

**1.6.26. Proposition.**

*Let $A$ be a set. Then $|\mathcal{P}(A)| > |A|$.*

*Proof.* We have an injection $A \to \mathcal{P}(A)$ by mapping $a \longmapsto \{a\}$, so $|A| \leq |\mathcal{P}(A)|$. Let $g : A \to \mathcal{P}(A)$ be any function. Consider the set $B = \{a \in A : a \notin g(a)\} \subset A$. We cannot have $g(a) = B$ for any $a \in A$. For if $g(a) = B$, then we get a contradiction as follows. If $a \in B$, then the definition of $B$ implies that $a \notin g(a) = B$, a contradiction; so $a \notin B$, which means that $a \in g(a) = B$, again a contradiction. It follows that $g$ cannot be surjective. In particular, no bijection $A \to \mathcal{P}(A)$ is possible. $\square$

And yet another application of the same idea; recall that we write $\{0,1\}^A$ to mean the set $\prod_A \{0,1\} = \{f : A \to \{0,1\}\}$.

**1.6.27. Proposition.**

*Let $A$ be a set. Then $|\{0,1\}^A| = |\mathcal{P}(A)|$.*

*Proof.* A smart way to prove this result is to notice that any $B \subset A$ can be identified with $1_B \in \prod_A \{0,1\}$, the characteristic function of $B$. The assignment $B \longmapsto 1_B$ is a bijection, and the result follows by Proposition 1.6.26. $\square$

A possible, and also common, approach is to first prove Proposition 1.6.27 (with Cantor's diagonal idea) and then deduce Proposition 1.6.26.

**1.6.28. Proposition.**

$|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.

*Proof.* As in Proposition 1.6.27 we have $|\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}|$. By thinking of each real number in its decimal representation we can construct an injection $\mathbb{R} \to \{0,1,\ldots,9\}^{\mathbb{N}}$; for instance we can map $\sum_{k=-n}^{\infty} d_k\, 10^{-k}$ to the function $g : \mathbb{N} \to \{0,1,\ldots,9\}$ given by

$$g(r) = \begin{cases} d_k, & r = 2k \\ d_{-k}, & r = 2k-1 \end{cases}$$

We have $|\{0,1,\ldots,9\}^{\mathbb{N}}| = |\{0,1\}^{\mathbb{N}}|$ (Exercise 1.6.7). Thus $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{N})|$.

Conversely, given $g \in \{0,1\}^{\mathbb{N}}$ we can map $g$ into $\mathbb{R}$ by

$$\gamma(g) = \sum_{k=1}^{\infty} g(k)\, 10^{-k}.$$

That is, we encode the sequence of zeros and ones as the decimal part of a real number in $(0,1)$. As $\gamma$ is injective, $|\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}| \leq |\mathbb{R}|$, and then Schröder–Bernstein gives us the equality $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. □

The question of whether there is an intermediate infinity between $|\mathbb{N}|$ and $|\mathbb{R}|$ was already considered by Cantor, and it was question number one in Hilbert's famous 23-problem list from 1900. The assumption that there is no such infinity in between became known as the *Continuum Hypothesis*. Gödel proved in 1940 that the negation of the continuum hypothesis cannot be proven from the standard Zermelo–Fraenkel set theory, even if one includes the Axiom of Choice. Cohen proved in 1963 that the continuum hypothesis can neither be proven from said axioms, thus showing that the continuum hypothesis is independent of the usual set theory and can be either affirmed or negated without contradiction. This implies that there exists some mysterious subset $X \subset \mathbb{R}$ such that we can assume $|X| = |\mathbb{R}|$ or $|X| < |\mathbb{R}|$ without contradicting any of the common set theory. As crazy as this sounds, it is worth noticing that for the vast majority of the subsets of $\mathbb{R}$, we cannot mention them explicitly in any possible way, and so the same goes for the bijections between such sets. The existence/description of said sets will depend on the axioms one uses and the functions available through those axioms. The conclusion is that $\mathbb{R}$ is so vast that the usual set theory cannot describe most of its subsets in any practical way.

We conclude the section with some more sophisticated results on cardinality. Many authors phrase them in terms of *cardinal arithmetic*, which in a sense they are, but we prefer the more straightforward formulation.

***1.6.29. Proposition. (Halving)***

*Let $A$ be an infinite set. Then there exist $A_1, A_2 \subset A$, disjoint, with $|A_1| = |A_2| = |A|$, and such that $A = A_1 \cup A_2$.*

*Proof.* Since $A$ is infinite, there exists $A_0 \subset A$ countable by Corollary 1.6.19. Let $\mathcal{F}$ be the collection of families $\{B_j\}$ of countable, pairwise disjoint subsets of $A$. We order $\mathcal{F}$ by inclusion. The existence of $A_0$ guarantees that $\mathcal{F}$ is nonempty. Given a chain in $\mathcal{F}$, its union will produce an upper bound; hence by Zorn's Lemma there exists a maximal family $\{B_j\}$ of countable pairwise disjoint subsets of $A$. The maximality guarantees that $A = \bigcup_j B_j$; for if $A \setminus B_j$ is finite we can put this finitely many points into some $B_j$, and if $A \setminus B_j$ is infinite then by Corollary 1.6.19 we are contradicting the maximality. As each $B_j$ is countable, we can write $B_j = C_j \cup D_j$ as a disjoint union of two countable sets. Let

$$A_1 = \bigcup_j C_j, \qquad A_2 = \bigcup_j D_j.$$

We have, as the unions are disjoint,

$$|A_1| = \left|\bigcup_j C_j\right| = \left|\bigcup_j B_j\right| = |A|,$$

and similarly $|A_2| = |A|$. And

$$A = \bigcup_j B_j = \bigcup_j (C_j \cup D_j) = \left(\bigcup_j C_j\right) \cup \left(\bigcup_j D_j\right) = A_1 \cup A_2. \qquad \square$$

***1.6.30. Lemma.***

*Let $A, B$ be nonempty sets with $A$ infinite and $B$ finite. Then $|A \cup B| = |A|$.*

*Proof.* Since $B \setminus A \subset B$ is finite, we may assume without loss of generality that $A \cap B = \varnothing$, since $A \cup B = A \cup (B \setminus A)$ with both sets disjoint. Write $B = \{b_1, \dots, b_m\}$. By Corollary 1.6.19 there exists $A_0 \subset A$ with $|A_0| = |\mathbb{N}|$. Denote the elements of $A_0$ by $\{a_1, a_2, \dots\}$. We define $g : A \cup B \to A$ by

$$g(a) = \begin{cases} a, & a \notin A_0 \cup B \\ a_j, & a = b_j \\ a_{j+m}, & a = a_j \end{cases}$$

Then $g$ is a bijection. $\square$

While it was very easy to give an explicit proof of Lemma 1.6.30, one can also use Corollary 1.6.18 and then in the disjoint case

$$|A \cup B| = |(A \cup B) \setminus B| = |A|.$$

**1.6.31. Definition.**

*Given sets $A, B$, their* disjoint union $A \sqcup B$ *is the set*

$$A \sqcup B = (A \times \{0\}) \cup (B \times \{1\}).$$

Of course the use of 0 and 1 is completely arbitrary. The point is that we consider elements of $A$ and $B$ as distinct, even if they are equal. The concrete formulation does not matter as we use this only for cardinality computations, where the actual sets do not matter, only their cardinality does.

***1.6.32. Proposition.***

*Let $A, B$ be nonempty sets with $A$ infinite and $|B| \leq |A|$. Then $|A \cup B| = |A|$.*

*Proof.* It is enough to show $|A \sqcup A| = |A|$, for then we would have

$$|A| \leq |A \cup B| \leq |A \sqcup A| = |A|,$$

and the equality follows from Schröeder–Bernstein (Theorem 1.6.13).

Let

$$\mathcal{F} = \{(X, f), \quad \text{where } f : X \times \{0,1\} \to X, \quad X \subset A \text{ and } f \text{ is bijective}\}.$$

We know that $\mathcal{F}$ is nonempty, because there exists $X \subset A$ countable (Corollary 1.6.19) and $|X \times \{0,1\}| = |X|$ by mapping, on $\mathbb{N}$, $(n,0) \longmapsto 2n-1$, $(n,1) \longmapsto 2n$ and using this bijection to obtain a corresponding bijection for $X$ through $|X| = |\mathbb{N}|$. We order $\mathcal{F}$ by saying that $(X,f) \leq (Y,g)$ if $X \subset Y$ and $f = g|_X$. A chain in $\mathcal{F}$ will be an increasing union $\{X_j, g_j)\}$. If we define $X = \bigcup X_j$ and define $g$ by $g|_{x_j} = g_j$, then $(X,g)$ is an upper bound for the chain; this requires showing that $g : X \times \{0,1\} \to X$ is bijective (Exercise 1.6.8). By Zorn's Lemma there exists a maximal element $(Z,h)$ on $\mathcal{F}$. If we show that $|Z| = |A|$, then we are done.

If $A \setminus Z$ is infinite, then it contains a countable infinite subset $Y$ by Corollary 1.6.19. Since $|Y| = |\mathbb{N}|$ we get, as in the previous paragraph, a bijection $k : Y \times \{0,1\} \to Y$. Then we can put $h$ and $k$ together to get a bijection $(Z \cup Y) \times \{0,1\} \to Z \cup Y$, contradicting the maximality of $Z$. Thus $A \setminus Z$ is finite. Then $Z$ is infinite, for otherwise $A = Z \cup (A \setminus Z))$ would be finite. Now, by Lemma 1.6.30

$$|Z| = |Z \cup (A \setminus Z)| = |A|.$$

□

***1.6.33. Proposition.***

*Let $A, B$ be nonempty sets with $A$ infinite and $|B| \leq |A|$. Then $|A \times B| = |A|$.*

*Proof.* It is enough to show that $|A \times A| = |A|$; for then we have

$$|A| \leq |A \times B| \leq |A \times A| = |A|,$$

and Schröeder–Bernstein (Theorem 1.6.13) gives the equalities.

Let

$$\mathcal{F} = \{(X, f), \text{ where } f : X \times X \to X, \quad X \subset A \text{ and } f \text{ is bijective}\}.$$

We know that $\mathcal{F}$ is nonempty, because there exists $X \subset A$ countable (Corollary 1.6.19) and $|X \times X| = |X|$ by Example 1.6.15. We order $\mathcal{F}$ by saying that

$(X, f) \leq (Y, g)$ if $X \subset Y$ and $f = g|_X$. A chain in $\mathcal{F}$ will be an increasing union $\{X_j, g_j)\}$. If we define $X = \bigcup X_j$ and define $g$ by $g|_{x_j} = g_j$, then $(X, g)$ is an upper bound for the chain; this requires showing that $g : X \times X \to X$ is bijective (Exercise 1.6.9). By Zorn's Lemma there exists a maximal element $(Z, h)$ on $\mathcal{F}$. If we show that $|Z| = |A|$, then we are done.

We know that $Z$ is infinite, because $\mathcal{F}$ contains a countable set. Suppose that $|Z| < |A|$. Then $|Z| < |A \setminus Z|$, for otherwise we would have $|A| = |(A \setminus Z) \cup Z| \leq |Z \times Z| = |Z|$, a contradiction. So there exists $Y \subset A \setminus Z$ with $|Y| = |Z|$; in particular, $|Y \times Y| = |Y|$. We have, using Proposition 1.6.32 and since $|Z \times Y| = |Y \times Z| = |Y \times Y|$,

$$|((Z \times Y) \cup (Y \times Z) \cup (Y \times Y)| = |Y \times Y|.$$

Thus

$$\begin{aligned}
|(Z \cup Y) \times (Z \cup Y)| &= |(Z \times Z) \cup (Z \times Y) \cup (Y \times Z) \cup (Y \times Y)| \\
&= |(Z \times Z) \sqcup \big[(Z \times Y) \cup (Y \times Z) \cup (Y \times Y)\big]| \\
&= |(Z \times Z) \sqcup (Y \times Y)| = |Z \cup Y|.
\end{aligned}$$

In the first three equalities above one can get the bijection to be the identity on $Z \times Z$. And then one can use $h$ in the last equality. This way we get a bijection $(Z \cup Y) \times (Z \cup Y) \to Z \cup Y$ that extends $h$, contradicting its maximality. It follows that $|Z| = |A|$. □

**_1.6.34. Corollary._**

*Let $A$ be an infinite set, and $\{B_a\}_{a \in A}$ a family of sets, indexed by $A$, with $|B_a| \leq |A|$ for all $a \in A$. Then*

$$\left| \bigcup_{a \in A} B_a \right| \leq |A|. \tag{1.9}$$

*Proof.* We have, using Proposition 1.6.33 in the last equality,

$$\left| \bigcup_{a \in A} B_a \right| \leq \left| \bigsqcup_{a \in A} B_a \right| \leq \left| \bigsqcup_{a \in A} A \right| = |A \times A| = |A|.$$ □

We cannot expect, in general, equality in (1.9); we could have for instance that the sets $B_a$ are all equal and finite, so it is even possibly for the union in (1.9) to be finite. We also mention that Corollary 1.6.34 contains as particular cases that the countable union of countable sets is countable (Example 1.6.17) and that a countable union of finite sets is countable.

### 1.6.1. Exercises.

(1.6.1) Show that equipotency is an equivalence relation.

(1.6.2) Let $n < m$ be positive integers. Show by induction that there is no bijection between $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$.

(1.6.3) Write an explicit bijection $\gamma : (0, 1) \to (0, 1) \cup (1, 2)$.

(1.6.4) Write $\gamma^{-1}$ explicitly for Example 1.6.5.

(1.6.5) Write in detail the proof of Proposition 1.6.27.

(1.6.6) Let $A$ be a set and $A = \bigcup_{j\in J} A_j$, with $\{A_j\}$ nonempty and pairwise disjoint. Fix sets $\{B_a\}_{a\in A}$. Show that

$$|\prod_{a\in A} B_a| = |\prod_{j\in J}\prod_{a\in A_j} B_a|.$$

(1.6.7) Let $A_1, A_2, \ldots$ be countable (finite or not), with $|A_n| \geq 2$ for all $n$. Show that

$$\left|\prod_{\mathbb{N}} A_n\right| = \left|\prod_{\mathbb{N}}\{0, 1\}\right|.$$

(1.6.8) Show that, in the proof of Proposition 1.6.32, $g : X \times \{0, 1\} \to X$ is a bijection.

(1.6.9) Show that, in the proof of Proposition 1.6.33, $g : X \times X \to X$ is a bijection.

## 1.7. Linear Algebra

Vector spaces are the most common objects in this book, so familiarity with the basics is crucial. While linear algebra is often presented first as the mathematics of solving linear systems, the point of view that we need here is that of vector spaces, linear dependence/independence, and bases.

**1.7.1. Definition.**

*A **field** is a triple $(\mathbb{F}, +, \times)$ where $\mathbb{F}$ is a set, $+$ is an operation that makes $(\mathbb{F}, +)$ an abelian group, namely:*

*(i)* $\alpha + \beta \in \mathbb{F}$ *for all* $\alpha, \beta \in \mathbb{F}$;

*(ii)* $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ *for all* $\alpha, \beta \in \mathbb{F}$ *(**associativity**);*

*(iii) there exists an element* $0 \in \mathbb{F}$*, called the **additive identity** such that* $\alpha + 0 = \alpha$ *for all* $\alpha \in \mathbb{F}$;

(iv) *for each $\alpha \in \mathbb{F}$, there exists $\beta \in \mathbb{F}$ such that $\alpha + \beta = 0$ (**additive inverse**);*

(v) *$\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in \mathbb{F}$ (**commutativity**).*

*And the multiplication $\times$, denoted by juxtaposition, satisfies*

(i) *$\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbb{F}$;*

(ii) *$\alpha(\beta\gamma) = (\alpha\beta)\gamma$ for all $\alpha, \beta, \gamma \in \mathbb{F}$ (**associativity**);*

(iii) *there exists an element $1 \in \mathbb{F} \setminus \{0\}$, called the **multiplicative identity** such that $1\alpha = \alpha$ for all $\alpha \in \mathbb{F}$;*

(iv) *for each $\alpha \in \mathbb{F} \setminus \{0\}$, there exists $\beta \in \mathbb{F}$ such that $\alpha\beta = 1$ (**multiplicative inverse**);*

(v) *$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ for all $\alpha, \beta, \gamma \in \mathbb{F}$ (**distributivity**).*

The elements 0 and 1 are unique (Exercise 1.1.7). The additive inverse and the multiplicative inverse, when they exist, are also unique (again Exercise 1.1.7). Because of this uniqueness we can denote the additive inverse of $\alpha$ by $-\alpha$ and the multiplicative inverse of $\alpha$ by $1/\alpha$ of $\alpha^{-1}$.

We have $0\alpha = 0$ for all $\alpha \in \mathbb{F}$ (Exercise 1.7.1). Also, $(-1)\alpha = -\alpha$, and $-(-\alpha) = \alpha$.

While there are lots of fields in existence, both finite and infinite, here the only fields that will be considered are $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, with their usual operations.

### 1.7.2. Definition.

*A **vector space** over a field $\mathbb{F}$ is an abelian group $(V, +)$ together with a multiplication $\mathbb{F} \times V \to V$ that satisfies, for $\alpha, \beta \in \mathbb{F}$ and $v, w \in V$,*

(i) *$\alpha(v + w) = \alpha v + \alpha w$;*

(ii) *$(\alpha + \beta)v = \alpha v + \beta v$;*

(iii) *$\alpha(\beta v) = (\alpha\beta)v$.*

*We say that $V$ is a vector space over $\mathbb{F}$. An element of $V$ is called a **vector**, and an element of $\mathbb{F}$ is called a **scalar**.*

In the particular case of a vector space over $\mathbb{R}$ or $\mathbb{C}$, we say that it is a real (complex) vector space. Any complex vector space is also a real vector space.

Though the notation is completely standard, it is important to recognize that $+$ means two completely different things here, namely the operation of the abelian group, and the addition in the field. The same happens with multiplication, we have the multiplication inside $\mathbb{F}$, and the multiplication of a scalar times a vector.

*1.7.3. Examples.* Here are some of the most common examples of vector spaces.

(a) For any $n \in \mathbb{N}$, the set $\mathbb{F}^n$, with pointwise operations, is a vector space over $\mathbb{F}$. In particular we have $\mathbb{R}^n$ and $\mathbb{C}^n$. We note that $\mathbb{C}^n$ is also a vector space over $\mathbb{R}$.

(b) The set $\mathbb{R}[x]$ of real polynomials is a real vector space.

(c) The set $\mathbb{C}[x]$ of complex polynomials is a complex vector space with the usual operations.

(d) The set $\{p \in \mathbb{C}[x] : \deg p \leq 5\}$ is a complex vector space.

(e) The set $C[0,1]$ of continuous functions $f : [0,1] \to \mathbb{C}$ is a complex vector space.

**1.7.4. Definition.**

*Given a vector space $V$, a subset $W \subset V$ is a* ***subspace*** *if $W$ is a vector space with the operations inherited from $V$. The smallest subspace of $V$ that contains $W$ is denoted by* $\operatorname{span} W$.

Here is some basic information about subspaces.

***1.7.5. Proposition.***

*Let $V$ be a vector space, and $W \subset V$ a subset. Then*

*(i)* *$W$ is a subspace if and only if $\alpha v + w \in W$ for all $\alpha \in \mathbb{F}$, $v, w \in W$.*

*(ii)* *The intersection of subspaces is a subspace; in particular, there exists a smallest subspace of $V$ that contains $W$.*

*(iii)* $\operatorname{span} W = \left\{ \sum_{j=1}^{m} \alpha_j w_j : \ m \in \mathbb{N},\ \alpha_1, \dots, \alpha_m \in \mathbb{F},\ w_1, \dots, w_m \in W \right\}.$

*Proof.* Exercise 1.7.3. □

**1.7.6. Definition.**

*Let $V$ be a vector space and $X \subset V$ a subset.*

*(i) We say that $X$ is **linearly independent** if whenever $\alpha_1 x_1 + \cdots + \alpha_m x_m = 0$ for $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and $x_1, \ldots, x_m \in X$, we have $\alpha_1 = \cdots = \alpha_m = 0$.*

*(ii) A subset that is not linearly independent is **linearly dependent**.*

*(iii) A **basis** is maximal linearly independent set.*

*(iv) The set $X$ is **generating** if $\operatorname{span} X = V$.*

*(v) An expression of the form $\displaystyle\sum_{j=1}^{m} \alpha_j w_j \;:\; m \in \mathbb{N},\ \alpha_1, \ldots, \alpha_m \in \mathbb{F},\ w_1, \ldots, w_m \in W$ is a **linear combination**.*

Here is some basic information about bases:

***1.7.7. Proposition.***

*Let $V$ be a vector space.*

*(i) A subset $X \subset V$ is linearly dependent if and only if there exists $x \in X$ such that $x \in \operatorname{span}(X \setminus \{x\})$.*

*(ii) A subset $X \subset V$ is a basis for $V$ if and only if $X$ is linearly independent and $V = \operatorname{span} X$.*

*(iii) Given a basis $X$, any element $v \in V$ can be written as a linear combination*

$$v = \sum_{j=1}^{m} \alpha_j x_j,$$

*where $\alpha_1, \ldots, \alpha_m \in \mathbb{F}$ and $x_1, \ldots, x_m \in X$, in a unique way.*

*(iv) Any two bases of $V$ have the same cardinality.*

*Proof.*

(i) Suppose that $X$ is linearly dependent. This means that there exist $\alpha_1, \ldots, \alpha_n$, not all zero, and $x_1, \ldots, x_n \in X$ such that $\sum_j \alpha_j x_j = 0$. Since addition is commutative, we may assume without loss of generality that $\alpha_1 \neq 0$. Then $x_1 = \sum_{j=2}^{n} \frac{-\alpha_j}{\alpha_1} x_j$. Conversely, if there exists $x_1 \in \operatorname{span}(X \setminus \{x_1\})$, then there exist $x_2, \ldots, x_n \in X$ and $\alpha_2, \ldots, \alpha_n \in \mathbb{F}$ such that $x_1 = \sum_{j=2}^{n} \alpha_j x_j$. Then $x_1 - \sum_{j=2}^{n} \alpha_j x_j = 0$ and the coefficient of $x_1$ is nonzero, so $X$ is linearly dependent.

(ii) Suppose that $X$ is a basis and that $\operatorname{span} X \subsetneq V$. Then there exists $v \in V$ such that $v \notin \operatorname{span} X$. Then necessarily $X \cup \{v\}$ is linearly independent, for if it were linearly dependent and there is a linear combination equal to zero where the coefficient of $v$ is zero, it would mean that $X$ is linearly dependent; and if there is a linear combination equal to zero where the coefficient of $v$ is nonzero, then $v \in \operatorname{span} X$. Thus $\operatorname{span} X = V$. Conversely, if $X$ is linearly independent and $V = \operatorname{span} X$, then there any element of $V$ is linearly dependent with $X$, and hence $X$ is maximal linearly independent.

(iii) This is a straight consequence of the fact that $V = \operatorname{span} X$.

(iv) Let $X, Y \subset V$ be two bases. Suppose that $X$ is finite. By writing each $x_j$ as in (1.10), we can use finitely many $y_k$ to span all of $V$; so $Y$ is finite. Thus if one basis is finite, all bases are finite. Write $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_m\}$. Since $x_1 \in \operatorname{span} Y$ and $x_1 \neq 0$, we can write $0 = x_1 - \sum_{j=1}^m \alpha_j y_j$, with at least one $\alpha_j$ nonzero; then $y_j \in \operatorname{span}\{\{x_1\} \cup (Y \setminus \{y_j\})$. By relabelling we may write this set as $Y_1 = \{x_1, y_2, \ldots, y_m\}$. Since $y_1 \in \operatorname{span} Y_1$, we still have $V = \operatorname{span} Y_1$. So now can do the same with $x_2$; and because $x_1, x_2$ are linearly independent, we necessarily have a nonzero coefficient for one of the remaining $y_j$. This allows to form—after relabelling the $y_j$ again—a set $Y_2 = \{x_1, x_2, y_3, \ldots, y_m\}$. Because $V = \operatorname{span} Y_2$ we may continue doing this, and there will always be $y_j$ available, for otherwise we would end up with $V = \operatorname{span}\{x_1, \ldots, x_m\}$ with $m < n$, contradicting the linear independence of $X$. Thus $n \leq m$, that $|X| \leq |Y|$. As the roles can be reversed, $|X| = |Y|$.

It remains to consider the case where $X, Y$ are infinite. By indexing with the set itself, we can write $X = \{x_j\}_{j \in J}$, $Y = \{y_k\}_{k \in K}$. For any $x_j$ we can write it in the basis $Y$ as

$$x_j = \sum_{k \in K_j} \alpha_{k,j} y_k, \tag{1.10}$$

where all the coefficients are nonzero. Then $X \subset \bigcup_{j \in J} \{y_k : k \in K_j\}$. Hence the union spans all of $V$, and this means that $Y = \bigcup_{j \in J} \{y_k : k \in K_j\}$ As each of the sets in the union is finite, we have that $|Y| \leq |J| = |X|$ by Corollary 1.6.34. As we can reverse the order, $|X| = |Y|$.

□

Bases are very non-unique. For starters, a permutation of a basis will again be a basis; and many linear combinations of the elements of a basis will again be a basis. For instance in $\mathbb{C}^n$ we have the $\{e_1, \ldots, e_n\}$, where $e_k(j) = \delta_{k,j}$. But if $A \in M_n(\mathbb{C})$ is an invertible matrix, the columns of $A$ (alternatively, the rows) form a basis of $\mathbb{C}^n$. A particular cases of this is that if you start with the canonical basis in $\mathbb{R}^2$, any independent rotations $f_1, f_2$ of $e_1$ and $e_2$ such that $f_1 \neq f_2$ makes another basis. In $\mathbb{C}[x]$ there is a canonical basis $\{1, x, x^2, \ldots\}$, but any family $\{p_n\}_{n=0}^{\infty}$ such that $\deg p_n = n$ is also a basis (Exercise 1.7.4); many many other bases are possible.

**1.7.8. Definition.**

*Let $V$ be a vector space and $X$ a basis for $V$. The **dimension** of $V$ is $|X|$, the cardinality of $X$.*

Dimension characterizes a vector space completely, as we will see. To discuss what "characterizes" means, we pay attention to the homomorphisms of linear spaces. These will be functions that preserve the vector space operations, and they are so common that they have their own particular name:

**1.7.9. Definition.**

*Let $V, W$ be vector spaces. A map $\phi : V \to W$ is **linear** if*

$$\phi(\alpha v_1 + v_2) = \alpha\phi(v_1) + \phi(v_2), \qquad \alpha \in \mathbb{F},\ v_1, v_2 \in V.$$

*If $\phi$ is also bijective, we say that it is an **isomorphism**, denoted $V \simeq W$.*

*The **kernel** and the **range** of $\phi$ are respectively the subspaces*

$$\ker\phi = \{v \in V : \ \phi(v) = 0\} \subset V, \qquad \operatorname{ran}\phi = \{\phi(v) : \ v \in V\} \subset W.$$

A straightforward computation shows that when $\phi$ is linear and bijective, $\phi^{-1}$ is also linear (Exercise 1.7.5).

***1.7.10. Proposition.***

*Let $V, W$ be vector spaces over the field $\mathbb{F}$. Then $V \simeq W$ if and only if $\dim V = \dim W$.*

*Proof.* Exercise 1.7.6. □

***1.7.11. Proposition.***

*Let $V$ be a vector space over $\mathbb{F}$ and let*

$$V^* = \{\varphi : V \to \mathbb{F}, \quad \textit{linear}\}.$$

*Show that*

*(i) $V^*$ is a vector space over $\mathbb{F}$ with $\dim V^* = \dim V$;*

*(ii)* *given a basis* $\{e_1, \dots, e_n\}$ *of* $V$ *there exists a basis* $\{e_1^*, \dots, e_n^*\}$ *of* $V^*$ *such that* $e_k^*(e_j) = \delta_{kj}$.

*Proof.* Exercise 1.7.7. □

The basis $\{e_1^*, \dots, e_n^*\}$ obtained in Proposition 1.7.11 is called the **dual basis** of $\{e_1, \dots, e_n\}$.

**1.7.12. Definition.**

*Let* $V$ *be a finite-dimensional vector space over the field* $\mathbb{F}$ *and* $\phi : V \to V$ *linear. The scalar* $\lambda \in \mathbb{F}$ *is an* ***eigenvalue*** *for* $\phi$ *if there exists nonzero* $v \in V$ *such that*

$$\phi(v) = \lambda v.$$

*The vector* $v$ *is said to be an* ***eigenvector*** *for* $\lambda$ *(and* $\phi$*).*

***1.7.13. Proposition.***

*Let* $V$ *be a finite-dimensional vector space and* $\phi : V \to V$ *linear. The following statements are equivalent:*

*(i)* $\lambda$ *is an eigenvalue for* $\phi$*;*

*(ii)* $\ker(\phi - \lambda I) \neq \{0\}$*;*

*(iii)* $\phi - \lambda I$ *is not invertible;*

*(iv)* $\det(\phi - \lambda I) = 0$.

*Proof.* Exercise 1.7.9. □

The condition $\det(\phi - \lambda I) = 0$ is particularly interesting because it provides a concrete (though not always practical) method to find eigenvalues. Indeed, the expression $p_\phi(\lambda) = \det(\phi - \lambda I)$ is a polynomial on $\lambda$, and so the eigenvalues of $\phi$ are precisely the roots of $p_\phi$. The polynomial $p_\phi$ is the **characteristic polynomial** of $\phi$.

When bases $E = \{e_1, \dots, e_n\}$ and $F = \{f_1, \dots, f_m\}$ are fixed for $V$ and $W$ respectively, a linear operator $\phi : V \to W$ can be seen canonically as a rectangular $m \times n$ matrix over $\mathbb{F}$. Indeed, for each $j$ there exist scalars $\{\alpha_{kj}\}$, $k = 1, \dots, m$, $j = 1, \dots, n$, such that

$$\phi(e_j) = \sum_{k=1}^{m} \alpha_{kj} \, f_k.$$

The notation is these matters is not very standardized; we will use $[\phi]_{E,F} = \alpha$ to denote the matrix representation of $\phi$ with respect to the bases $E$ and $F$.

The usual addition (entrywise) and multiplication ("row times column") of matrices correspond to the addition and composition of linear maps, respectively (Exercises 1.7.10 and 1.7.11). Namely, if $\psi_1, \psi_2 : V \to W$ are linear and $E, F$ are bases for $V, W$ respectively, then

$$[\phi_1 + \phi_2]_{E,F} = [\phi_1]_{E,F} + [\phi_2]_{E,F}.$$

And if $\phi : V \to W$ and $\psi : W \to Z$ with respect to bases $E, F, G$ of $V, W, Z$ respectively, then

$$[\psi \circ \phi]_{E,G} = [\psi]_{F,G}[\phi]_{E,F}.$$

When $\mathbb{F}$ is algebraically closed (or, alternatively, if we work on the algebraic closure of $\mathbb{F}$), the characteristic polynomial can be written $p_\phi(x) = \alpha(x - \lambda_1) \cdots (x - \lambda_n)$, where $\lambda_1, \ldots, \lambda_n$ are its roots; and these are precisely the eigenvalues of $\phi$, with possible repetitions, that we count as (algebraic) "multiplicities".

In general, for most bases the matrix form of the operator $\phi$ will not expose the eigenvalues explicitly, but it may do for some well-chosen bases. In particular, every operator on a finite-dimensional space is *triangularizable*:

***1.7.14. Proposition. (Triangularization)***

*Let $V$ be a finite-dimensional space over an algebraically closed field $\mathbb{F}$, and $\phi : V \to V$ linear. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\phi$, counting multiplicities, then there exists a basis $e_1, \ldots, e_n$ of $V$ and scalars $\{\alpha_{kj}\}$ such that*

$$\phi(e_k) = \lambda_k e_k + \sum_{j=1}^{k-1} \alpha_{kj} e_j. \tag{1.11}$$

*When $V$ is an inner product space, the basis can be chosen to be orthonormal.*

*Proof.* We proceed by induction on $\dim V$. If $\dim V = 1$, then by definition of eigenvalue there exists nonzero $v_1 \in V$ with $Tv_1 = \lambda_1 v_1$. Now assume as inductive hypothesis that if $\dim V < n$, then (1.11) holds for adequate vectors and scalars. When $\dim V = n$, we have that $\ker(\phi - \lambda_n I) \neq \{0\}$. Let $W = (\phi - \lambda_n I)V$. Since $\phi - \lambda_n I$ is not injective, $\dim W < \dim V$ (see Exercise 1.7.8). Also, if $w \in W$, we can write $w = (\phi - \lambda_n I)v = \phi(v) - \lambda_n v$ for some $v \in V$. Then

$$\phi(w) = \phi(\phi(v)) - \lambda_n \phi(v) = (\phi - \lambda_n I)\phi(v) \in W.$$

So $W$ is invariant for $\phi$ and the inductive hypothesis applies. This means that there exists a (orthonormal, when an inner product is present) basis

$f_1, \dots, f_m$ of $W$ and scalars $\{\gamma_{kj}\}$ such that

$$\phi(f_k) = \mu_k f_k + \sum_{j=1}^{k-1} \gamma_{kj} f_j,$$

where $\mu_1, \dots, \mu_m$ are eigenvalues of $\phi|_W$. Since each $\mu_k$ is an eigenvalue for $\phi$, we necessarily have $\{\mu_1, \dots, \mu_m\} \subset \{\lambda_1, \dots, \lambda_{n-1}\}$.

Let us complete the linearly independent set $\{f_1, \dots, f_m\}$ to a (orthonormal, when an inner product is present) basis $\{f_1, \dots, f_m, g_1, \dots, g_r\}$. We have, since $(\phi - \lambda_1)g_t \in W$,

$$\phi(g_t) = (\phi - \lambda_n I)g_t + \lambda_n g_t = \sum_{\ell=1}^{m} \beta_{\ell,t} f_\ell + \lambda_n g_t.$$

Now we can take $e_j = f_j$ for $j = 1, \dots, m$, and $e_{j+r} = g_j$ for $j = 1, \dots, r$; and $\alpha_{kj} = \gamma_{kj}$ for $k = 1, \dots, m$ and $j = 1, \dots, k-1$, $\alpha_{kj} = \beta_{k-m,j}$ for $k = m+1, \dots m+r$ and $j = 1, \dots, k-m-1$.

With respect to the basis $E$ we have (after possibly reordering the basis)

$$[\phi]_{E,E} = \begin{bmatrix} \mu_1 & \alpha_{21} & \alpha_{22} & * & * \\ 0 & \mu_2 & \alpha_{31} & * & \vdots \\ 0 & 0 & \mu_3 & * & \vdots \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{bmatrix}.$$

The characteristic polynomial is $p_\phi(t) = (t - \lambda_n)^{n-m} \prod_{j=1}^{m} (t - \mu_j)$, and thus

$$\{\mu_1, \dots, \mu_m\} = \{\lambda_1, \dots, \lambda_m\} \qquad \square$$

.

So every operator acting on a finite-dimensional space is triangularizable. The requirement that $\mathbb{F}$ is algebraically closed is not too onerous, for any field admits an algebraic closure.

There is a more specialized form of triangularization, the **Jordan Form** of an operator.

**1.7.15. Definition.**

*Given $\lambda \in \mathbb{C}$ and $k \in \mathbb{N}$, the **Jordan block** $J_k(\lambda)$ is the $k \times k$ matrix*

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

*A matrix $A \in M_n(\mathbb{C})$ is in **Jordan form** if $A$ is a block matrix*

$$A = \begin{bmatrix} J_{k_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{k_r}(\lambda_r) \end{bmatrix}.$$

**1.7.16. Proposition. (Jordan Form)**

*Let $A \in \mathcal{B}(\mathbb{C}^n)$. Then there exists a basis $G = \{g_1, \dots, g_n\}$ of $\mathbb{C}^n$ such that the matrix of $A$ with respect to $G$ is in Jordan form. The Jordan form of $A$ is unique up to permutation of Jordan blocks.*

Necessarily, the diagonal of the Jordan form of $A$ contains the eigenvalues of $A$, counting multiplicities.

**1.7.17. Definition.**

*Given $A \in \mathcal{B}(\mathbb{C}^n)$ and $\lambda \in \sigma(A)$, the **geometric multiplicity** of $\lambda$ is the number of Jordan blocks corresponding to $\lambda$, and the **algebraic multiplicity** of $\lambda$ is the sum of the $k_j$ such that $\lambda_j = \lambda$.*

The geometric multiplicity of $\lambda$ agrees with $\dim\ker(A - \lambda I)$, and the algebraic multiplicity agrees with the multiplicity of $\lambda$ as a root of the characteristic polynomial of $A$.

### 1.7.1. Exercises.

(1.7.1) Let $\mathbb{F}$ be a field. Show that $0\alpha = 0$ for all $\alpha \in \mathbb{F}$.

(1.7.2) Let $\mathbb{F}$ be a field, $\alpha, \beta \in \mathbb{F}$. Show that

(i) $(-\alpha)\beta = \alpha(-\beta) = -(\alpha\beta)$;

(ii) $-(-\alpha) = \alpha$;

(iii) $(-1)\alpha = -\alpha$.

(1.7.3) Prove Proposition 1.7.5.

(1.7.4) Let $\{p_n\}_{n=0}^{\infty} \subset \mathbb{F}[x]$ such that $\deg p_n = n$ for all $n$. Show that $\{p_n\}$ is a basis for $\mathbb{F}[x]$.

(1.7.5) Show that if $V, W$ are vector spaces and $\phi : V \to W$ is bijective and linear, then $\phi^{-1}$ is linear.

(1.7.6) Prove Proposition 1.7.10.

(1.7.7) Prove Proposition 1.7.11.

(1.7.8) Let $V, W$ be a finite-dimensional vector spaces with $\dim W = \dim V$ and $\phi : V \to V$ linear. Show that $\phi$ is injective if and only if it is surjective.

(1.7.9) Prove Proposition 1.7.13.

(1.7.10) Let $V, W$ be vector spaces over $\mathbb{F}$ and $\phi_1, \phi_2 : V \to W$ linear. Fix bases $E, F$ for $V$ and $W$ respectively, and show that

$$[\phi_1 + \phi_2]_{E,F} = [\phi_1]_{E,F} + [\phi_2]_{E,F}.$$

(1.7.11) Let $V, W, Z$ be vectors spaces over $\mathbb{F}$ and $\phi : V \to W$, $\psi : W \to Z$ linear maps. Fix bases $\{e_1, \dots, e_n\}$ for $V$, $\{f_1, \dots, f_m\}$ for $W$, and $\{g_1, \dots, g_p\}$ for $Z$. Show that

$$[\psi \circ \phi]_{E,G} = [\psi]_{F,G}[\phi]_{E,F}.$$

(1.7.12) Let $V$ be a finite-dimensional vector space and $\phi : V \to V$ linear. Let $\lambda_1, \dots, \lambda_n$ be distinct eigenvalues for $\phi$. Show that if $v_1, \dots, v_n$ are eigenvectors for $\lambda_1, \dots, \lambda_n$ respectively, then $v_1, \dots, v_n$ are linearly independent.

## 1.8. Basic Point Set Topology

We begin with a result that will be needed later. It shows that open sets in $\mathbb{R}$ cannot be too crazy.

**1.8.1. Proposition.**

*Let $V \subset \mathbb{R}$ be open. Then there exists a pairwise disjoint family $\{I_n\}_{n\in N}$, with $N \subset \mathbb{N}$, of open intervals such that $V = \bigcup_n I_n$.*

*Proof.* Given $x \in V$, define

$$a_x = \inf\{t \in \mathbb{R} : (t, x) \subset V\}, \qquad b_x = \sup\{s \in \mathbb{R} : (x, s) \subset V\}.$$

As $V$ is open both numbers are well defined, provided that we allow $-\infty$ for $a_x$ and $+\infty$ for $b_x$. Also from $V$ open we obtain that $a_x \notin V$ (if it were, $a_x - \delta \in V$ for some $\delta > 0$) and $b_x \notin V$. Also, $(a_x, b_x) \subset V$; this follows from $(a_x, x) \subset V$, $(x, b_x) \subset V$ and $x \in V$ with $V$ open.

Define a relation on $V$, by $x \sim y$ if $x \in (a_y, b_y)$. In such case, since $x \in (a_y, b_y) \subset V$, we immediately have that $(a_y, x) \subset V$, and so $a_x \leq a_y$. Similarly, $(x, b_y) \subset V$ and so $b_y \leq b_x$. Thus $(a_y, b_y) \subset (a_x, b_x)$; this shows that $y \in (a_x, b_x)$, implying that $\sim$ is symmetric. It also shows that $(a_x, b_x) = (a_y, b_y)$ after exchanging roles. Now transitivity follows directly. In fact, every time that a relation is defined in terms of equality, it has to be an equivalence relation. So $V$ is partitioned into its equivalence classes, which are open disjoint intervals. There are several ways to verify that there can only be at most countably many such intervals. For instance, for each class take a rational representative; then the set of representatives is either finite or countably infinite. □

*1.8.2. Remark.* Here is a more conceptual description of the proof of Proposition 1.8.1. Given $x \in V$, there exists an open interval $I \subset V$ that contains it. Let $I_x$ be the union of all open intervals that contain $x$ and lie inside $V$. A union of open intervals with a common point is itself an open interval. It is routine to see that $I_x$ and $I_y$ are either equal or disjoint. So $V$ is a union of these intervals. As each contains a rational number, there can be at most countably many.

### 1.8.1. Metric Spaces.

**1.8.3. Definition.**

*A **metric space** is a pair $(X, d)$, where $X$ is a set and $d : X \times X \to [0, \infty)$ is a **distance**:*

*(i) $d(x, y) = 0$ if and only if $x = y$;*

*(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$ (Symmetry);*

*(iii) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$ (Triangle Inequality).*

*We write $\lim_{n\to\infty} x_n = x$ if $\lim_{n\to\infty} d(x_n, x) = 0$. The **ball** of radius $r$ around $x$ is the set*

$$B_r^X(a) = \{x \in X : \ d(x, a) < r\}.$$

*When context is clear, we may write $B_r(a)$, without the $X$.*

A useful inequality that will be used repeatedly is the **reverse triangle inequality** (Exercise 1.8.24):

$$|d(x, y) - d(y, z)| \leq d(x, z), \qquad\qquad x, y, z \in X.$$

The convergence $x_n \to x$ can be stated as saying that for every $\varepsilon > 0$ there exists $n_0$ such that $x_n \in B_\varepsilon(x)$ for all $n \geq n_0$. Or, in other words, $x_n \to x$ if and only if for every ball that contains $x$ eventually the sequence falls inside the ball.

We say that a sequence $\{x_n\} \subset X$ is **Cauchy** if for all $\varepsilon > 0$ there exists $n_0$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq n_0$. Every convergent sequence is Cauchy. When every Cauchy sequence is convergent, we say that $X$ is **complete**. With the usual notion of convergence (that is, with the distance $d(x, y) = |x - y|$) $\mathbb{R}$ is complete while $\mathbb{Q}$ is not complete.

A set $E \subset X$ is **open** if for each $e \in E$ there exists a ball $B_r(a)$ such that $e \in B_r(a) \subset E$. The triangle inequality guarantees that balls are open (Exercise 1.8.1). A set $E \subset X$ is **closed** if its complement $X \setminus E$ is open. It is follows directly from the definitions that arbitrary unions of open sets are open, and that arbitrary intersections of closed sets are closed.

Given metric spaces $X, Y$, a function $f : X \to Y$ is **continuous** at a point $x_0 \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d_X(x, x_0) < \delta$ implies $d_Y(f(x), f(x_0)) < \varepsilon$. This is equivalent to saying that given any ball $B_1 \subset Y$ that contains $f(x_0)$, there is a ball $B_0 \subset X$ around $x_0$ such that $f(B_0) \subset B_1$. Going more abstract, but still equivalent, we can say that for every ball $B \subset Y$ with $f(x_0) \in B$, the preimage $f^{-1}(B)$ is open. This means that $f$ is continuous everywhere in $X$ if and only if $f^{-1}(E)$ is open for all open sets $E \subset Y$. We denote by $C(X)$ the set of complex-valued continuous functions on $X$; in particular, $C[0, 1]$ denotes the complex-valued continuous functions on $[0, 1]$.

Many natural metric spaces are not complete. For instance $\mathbb{Q}$ with the usual metric, but we will find others as we go along. It turns out that any metric space can be completed. Properly, a **completion** of a metric space $X$ is a metric space $\tilde{X}$ such that

(i) $\tilde{X}$ is complete;

(ii) there is an isometric injection (that is, an **embedding**) $\gamma : X \to \tilde{X}$, i.e. $\gamma$ is injective and
$$d_{\tilde{X}}(\gamma(x), \gamma(y)) = d_X(x, y), \qquad x, y \in X;$$

(iii) $\gamma(X)$ is dense in $\tilde{X}$.

Intuitively, we can somehow think of $X$ inside $\tilde{X}$ while keeping its metric space structure, and every Cauchy sequence in $X$ is now convergent in $\tilde{X}$. We have

***1.8.4. Proposition.***

*Let $X$ be a metric space with metric $d$. Then there exists a completion $(\tilde{X}, \gamma)$. The completion is unique in the sense that between any two completions $(\tilde{X}_1, \gamma_1)$ and $(\tilde{X}_2, \gamma_2)$ there is a bijective isometry $\beta : \tilde{X}_1 \to \tilde{X}_2$ with $\beta \circ \gamma_1 = \gamma_2$.*

*Proof.* We provide a more abstract, functional-analytical, and economical proof, while the more natural, intuitive proof is highlighted in Exercise 1.8.26.

Consider $C_b(X, \mathbb{R})$, the bounded continuous real-valued functions on $X$, with the **uniform metric**
$$d'(f, g) = \sup\{|f(x) - g(x)| : x \in X\}.$$

Fix $b \in X$. We define $\rho : X \to C_b(X, \mathbb{R})$ by $\rho(x) = \hat{x}$, where $\hat{x}(y) = d(x, y) - d(y, b)$. Then $\hat{x}$ is bounded, since

$$|\hat{x}(y)| = |d(x, y) - d(y, b| \leq d(x, b), \qquad y \in X.$$

And $\hat{x}$ is continuous since

$$\begin{aligned}|\hat{x}(y) - \hat{x}(z)| &= |d(x, y) - d(y, b) - (d(x, z) - d(z, b))| \\ &\leq |d(x, y) - d(x, z)| + |d(y, b) - d(z, b)| \leq 2d(y, z).\end{aligned}$$

The function $\rho$ is injective: if $\rho(x) = \rho(z)$, then $d(x, y) - d(y, b) = d(z, y) - d(y, b)$ for all $y \in X$; hence $d(x, y) = d(z, y)$ and in particular,

$$0 = d(x, x) = d(z, x),$$

so $d(z, x) = 0$ and thus $z = x$. Also, $\rho$ is isometric. Indeed, once again using the reverse triangle inequality,

$$\begin{aligned}d'(\hat{x}, \hat{z}) &= \sup_{y \in X} \{|d(x, y) - d(y, b) - (d(z, y) - d(y, b))|\} \\ &= \sup_{y \in X} \{|d(x, y) - d(z, y)|\} \leq d(x, z).\end{aligned}$$

The supremum is achieved when $y = z$, thus $d'(\hat{x}, \hat{z}) = d(x, z)$.

We define $\bar{X}$ as the closure of $\rho(X)$ in $C_b(X, \mathbb{R})$. The density of $\rho(X)$ in $\bar{X}$ then occurs by definition, and all we need to do is show that $\bar{X}$ is complete. For this it is enough to show that $C_b(X, \mathbb{R})$ is complete, as a closed subset of a complete metric space will be complete (Exercise 1.8.25). Let $(g_n)$ be Cauchy in $C_b(X, \mathbb{R})$. For each $x \in X$, we have $|g_m(x) - g_n(x)| \leq d'(g_m, g_n)$, and so the sequence $\{g_n(x)\}$ is Cauchy in $\mathbb{R}$. Thus we can define the pointwise limit $g(x) = \lim_n g_n(x)$. We will see that the convergence is uniform, giving us that $g$ is continuous (Exercise 1.8.16), so $g \in C_b(X, \mathbb{R})$. Indeed, given $\varepsilon > 0$ there exists $n_0$ such that $d'(g_m, g_n) < \varepsilon$ whenever $m, n > n_0$. Then, for $m, n > n_0$,

$$|g(x) - g_n(x)| \leq |g(x) - g_m(x)| + |g_m(x) - g_n(x)| \leq |g(x) - g_m(x)| + \varepsilon.$$

This works for any $m$, so taking limit as $m \to \infty$ we have

$$|g(x) - g_n(x)| < \varepsilon, \qquad x \in X, n > n_0,$$

showing that $d'(g, g_n) < \varepsilon$ and so $g_n \to g$ uniformly. As mentioned above, being a uniform limit of continuous functions, $g$ is continuous. And $g$ is bounded, since there exists $n$ such that $d'(g, g_n) < 1$ and $c >$ with $|g_n(x)| \leq c$ for all $x$; then

$$|g(x)| \leq |g(x) - g_n(x)| + |g_n(x)| \leq 1 + c,$$

and then $g \in C_b(X, \mathbb{R})$.

It remains to show the uniqueness of the completion. Let $X_1, X_2$ be completions of $X$ with isometric embeddings $\gamma_1 : X \to X_1$, $\gamma_2 : X \to X_2$.

Let $\rho : \gamma_1(X) \to \gamma_2(X)$ be given by $\rho \circ \gamma_1 = \gamma_2$. This well defined, since $\gamma_1$ is injective. Also,

$$d_2(\rho(\gamma_1(x)), \rho(\gamma_1(y)) = d_2(\gamma_2(x), \gamma_2(y)) = d(x, y) = d_1(\gamma_1(x), \gamma_1(y)),$$

so $\rho$ is isometric. For any $a \in X_1$, there exists a sequence $(x_n) \subset X$ with $\gamma_1(x_n) \to a$. As the sequence $\{\gamma_1(x_n)\}$ is Cauchy and $\rho$ is isometric, the sequence $\{\rho(\gamma_1(x_n))\}$ is also Cauchy, so it has a limit that we denote $\rho(a) \in X_2$. This is well-defined, because if $\gamma_1(y_n) \to a \in X_1$ and $\gamma_2(y_n) \to b \in X_2$, then

$$\begin{aligned} d_2(\rho(a), b) &= \lim_n d_2(\rho(\gamma_1(x_n)), \rho(\gamma_1(y_n))) \\ &= \lim_n d_1(\gamma_1(x_n), \gamma_1(y_n)) = d_1(a, a) = 0. \end{aligned}$$

As $\rho$ is isometric and maps a dense subset to a dense subset, it is bijective and isometric (Exercise 1.8.28). □

A metric space $X$ is **separable** if there exists $X_0 \subset X$, countable and dense.

***1.8.5. Proposition.***

*Let $X$ be a separable metric space. If $Y \subset X$, then $Y$ is separable.*

*Proof.* By hypothesis there exists $X_0 \subset X$, countable and dense. Given any $m \in \mathbb{N}$ and $x \in Y$, there exists $x_0 \in X_0$ such that $x_0 \in B_{1/m}(x)$; we can read this statement as saying that $x \in B_{1/m}(x_0)$. Hence

$$Y \subset \bigcup_{x_0 \in X_0} \bigcup_m B_{1/m}(x_0).$$

Form $Y_0$ by taking one element from $Y \cap B_{1/m}(x_0)$, $m \in \mathbb{N}$, $x_0 \in X_0$. Then $Y_0$ is countable by construction, and if $y \in Y$ and $\varepsilon > 0$, choose $m > \frac{2}{\varepsilon}$ and $x_0 \in X_0$ such that $d(y, x_0) < \varepsilon/2$. By construction there exists $y_0 \in Y_0 \cap B_{1/m}(x_0)$, and

$$d(y, y_0) \leq d(y, x_0) + d(x_0, y_0) < \frac{\varepsilon}{2} + \frac{1}{m} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

So $y \in \overline{Y_0}$. □

**1.8.2. Topological Spaces.** If one looks carefully at the definitions regarding continuity, limits, and the like, they are phrased in terms of open subsets. As far as the definitions go, it doesn't matter much what the actual sets are as long as they behave in a certain way. The idea of a topological space came out of this point of view. Instead of defining the open sets in terms of some property—distance, in the case of metric spaces—they are given a priori; of course, with some properties that guarantee that they behave they we one expects.

Recall that given a set $X$, its **power set** is the set $\mathcal{P}(X)$ of all subsets of $X$. A **topology** on $X$ is a family $\mathcal{T} \subset \mathcal{P}(X)$ such that

(1) $\varnothing \in \mathcal{T}$, $X \in \mathcal{T}$;
(2) if $A_j \in \mathcal{T}$ for all $j \in J$, then $\bigcup_j A_j \in \mathcal{T}$;
(3) if $A_1, A_2 \in \mathcal{T}$, then $A_1 \cap A_2 \in \mathcal{T}$.

This definition retains the properties that make balls and open sets useful in a metric space, but now with no reference to a metric. So we call the sets in $\mathcal{T}$ open, and their complements closed. A metric space $X$ is a topological space if we take $\mathcal{T}$ to be the collection of all open subsets of $X$.

The definition of topological space feels bare, and it is. We will see soon that very pathological topological spaces exist, in the sense of having poor separation properties, see below.

An intersection of topologies is a topology, and $\mathcal{P}(X)$ is always a topology. This implies that given a family $\mathcal{T}_0 \subset \mathcal{P}(X)$, there is always a smallest topology $\operatorname{Top}(\mathcal{T}_0)$ that contains $\mathcal{T}_0$. We say that $\operatorname{Top}(\mathcal{T}_0)$ is the **topology generated** by $\mathcal{T}_0$.

Given two topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ on a set $X$, we say that $\mathcal{T}_1$ is **weaker** (or **coarser**) than $\mathcal{T}_2$ if $\mathcal{T}_1 \subset \mathcal{T}_2$. In this same situation we say that $\mathcal{T}_2$ is **stronger** (**finer**) than $\mathcal{T}_1$. As can be easily seen from the definitions below, a weaker topology has less continuous functions and more compact sets than a stronger topology.

A topological space $X$ is **Hausdorff** if for any $x, y \in X$ there exist $V, W \subset X$, open, with $x \in V$, $y \in W$, and $V \cap W = \varnothing$. Every metric space is Hausdorff, but not every topological space is Hausdorff. For a pathological example consider the set $X = \{1, 2, 3\}$ with the topology

$$\mathcal{T} = \{\varnothing, X, \{1\}, \{2, 3\}\}.$$

The points 2 and 3 cannot be separated. In particular, the constant sequence $\{2\}$ converges to 2 (no surprise here), but it also converges to 3. Indeed, any open neighbourhood $V$ of 3 is either $X$ or $\{2, 3\}$, so $2 \in V$ regardless. Non-Hausdorff topological spaces are not pretty.

There is a hierarchy of separation axioms for topological spaces, often labelled $T_0, T_1, \ldots, T_6$, with some fractional types too. Here Hausdorff is $T_2$. We will also name $T_4$, the so-called normal spaces. A topological space $X$ is **normal** if given any two disjoint closed sets $C_1, C_2$, there exist disjoint open sets $V_1 \supset C_1$ and $V_2 \supset C_2$. Every metric space is normal (Exercise 1.8.5), and every compact Hausdorff space is normal (Exercise 2.6.1).

We recall that, formally, a sequence $\{a_n\}$ in $X$ is a function $a : \mathbb{N} \to X$. What makes a sequence a sequence is the fact that $\mathbb{N}$ is countable and, equally importantly, that $\mathbb{N}$ is well-ordered. In a metric space sequences can be used to approximate limit points, and to establish continuity of functions. To deal with non-metric topological spaces, it is essential to use nets rather than sequences. A **net** in $X$ is a function $a : J \to X$, where $J$ is a **directed** set; this means that $J$

has a (not necessarily total) preorder (that is, a reflexive and transitive relation) and that every finite subset of $J$ has an upper bound in $J$. The prototypical example of a directed set that is not totally ordered is that of the finite subsets of a set, ordered by inclusion.

Given a net $\{a_j\}_{j\in J}$ in a topological space $X$, we say that $\lim_j a_j = a$ if for every open $V \subset X$ with $a \in V$, there exists $j_0 \in J$ such that $a_j \in V$ whenever $j \geq j_0$. The index set $J$ should always be clear from context.

We will see soon (Proposition 1.8.7) that a topology is determined by convergence.

The reason sequences—instead of nets—suffice in a metric space is that if any point $x$ is in the boundary of a set $K$, then one can consider the sequence of balls $B_{1/n}(x)$ so that for any ball $B_r(x)$ there is an $n$ such that $B_{1/n}(x) \subset B_r(x)$; in particular, for each $n$ there exists $x_n \in K$ such that $x_n \in B_{1/n}(x)$. Then the sequence $\{x_n\}$ converges to $x$. In a topological space it might be impossible to do something similar in a countable way, hence the need for nets. Now if $x$ is in the boundary of $K$, for each neighbourhood $W$ of $x$ we can find $x_W \in K$ with $x_W \in W$. Then the net $\{x_W\}$ converges to $x$, but there might not be a sequence of elements of $K$ that converges to $x$.

We recall that when every Cauchy net is convergent, the space is **complete**.

In order to obtain a useful definition, it is not a good idea to define subnet by just mimicking what we do for a subsequence; that is, just restricting indices is not good enough. Instead, a **subnet** of a net $\{x_j\}_{j\in J}$ is $\{x_{h(k)}\}_{k\in K}$, where $K$ is a directed set (no relation with $J$) and $h : K \to J$ is monotone and cofinal. Here **cofinal** means that for every $j \in J$ there exists $k \in K$ with $h(k) \geq j$. This is the definition that is required for "subnet" to behave topologically like the analog of "subsequence".

Since unions of open sets are open and closed sets are complements of open, it follows that intersections of closed sets are always closed. So, given any subset $E$ of $X$, the intersection of all closed subsets of $X$ that contain $E$ is the smallest closed set that contains $E$; we call it the **closure** of $E$, denoted $\overline{E}$.

Sometimes some topological manipulations are easier when dealing with some non-open sets (like closed intervals in $\mathbb{R}$). A **neighbourhood** $N$ of $x \in X$ is a subset of $X$ such that $x \in N$ and there exists $V \subset X$, open, with $x \in V \subset N$. The usefulness of neighbourhoods comes from allowing us to discuss the topology via easier to handle sets. Given $A \subset X$, a point $b \in X$ is a **cluster point** (or **accumulation point**) for $A$ if for every neighbourhood $N$ of $b$ we have $(N \setminus \{b\}) \cap A \neq \varnothing$. In other words, $b$ is a cluster point if it is a limit of points in $A$, where the constant net $\{b\}$ is not allowed. A point $a \in X$ is a **boundary point** for $A$ if for any neighbourhood $N$ of $a$, $N \cap A \neq \varnothing$ and $N \cap A^c \neq \varnothing$.

To help with picturing the difference between a cluster point and a boundary point, let $A$ be the open disk of radius 1, centered at the origin; then the set of cluster points of $A$ is the closed disk of radius 1 centered at the origin, while the boundary points are the circle of radius 1, centered at the origin. In the set $\{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}\} \subset \mathbb{R}$, the only cluster point is 0, while all its points are boundary points.

If we denote by $C(A)$ the set of cluster points of $A$, and $\partial A$ denotes the set of boundary points of $A$, we have

**1.8.6. Proposition.**

*Let $X$ be a topological space, $A, B \subset X$. Then*

*(i)* $A$ *is closed if and only if* $C(A) \subset A$.

*(ii)* $A$ *is closed if and only if* $\partial A \subset A$.

*(iii)* $\overline{A} = A \cup C(A)$.

*(iv)* $\overline{A} = A \cup \partial A$

*(v)* $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

*(vi)* $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$; *the inclusion may be strict.*

The **interior** of $A \subset X$ is the set $\operatorname{Int} A = \bigcup\{V \subset X : \text{ open}, V \subset A\}$. Then

$$(\operatorname{Int} A) \cap \partial A = \varnothing, \qquad (\operatorname{Int} A) \cup \partial A = \overline{A}.$$

**1.8.7. Proposition.**

*If $X$ has topologies $\mathcal{T}_1$ and $\mathcal{T}_2$, then the following statements are equivalent:*

*(i)* $\mathcal{T}_1 = \mathcal{T}_2$

*(ii) a net converges to $x$ in $\mathcal{T}_1$ if and only if it converges to $x$ in $\mathcal{T}_2$.*

*Proof.* (i) $\Longrightarrow$ (ii) trivial.

(ii) $\Longrightarrow$ (i) Let $V \in \mathcal{T}_1$. We will use the notation $C_1$ and $C_2$ to denote the set of cluster points of a set in each topology. Given $x \in C_1(V^c)$, by hypothesis we get that $x \in C_2(V^c)$; it follows that $C_1(V^c) = C_2(V^c)$. As $V^c$ is closed, by Proposition 1.8.6 we have that $C_1(V^c) \subset V^c$, and hence $C_2(V^c) \subset V^c$; again by Proposition 1.8.6, $V^c$ is closed in $\mathcal{T}_2$, and so $V \in \mathcal{T}_2$. As the roles can be reversed, $\mathcal{T}_1 = \mathcal{T}_2$. $\square$

A very common application of the lim sup, that we will use often, is the following:

**1.8.8. Proposition. (The Limsup Routine)**

*Let $X$ be a topological space, $\{x_j\} \subset X$ a net, and $f : X \to [0, \infty)$ a function. If*

$$\limsup_j |f(x_j)| < \varepsilon$$

*for all $\varepsilon > 0$, then $\lim_j f(x_j)$ exists and is equal to $0$.*

*Proof.* As $\varepsilon > 0$ is arbitrary, the inequality $\limsup_j |f(x_j)| < \varepsilon$ implies that $\limsup_j |f(x_j)| = 0$. Then

$$0 \leq \liminf_j |f(x_j)| \leq \limsup_j |f(x_j)| = 0.$$

Since the liminf and the limsup agree and are equal to 0, the limit exists and it is equal to 0 by Exercise 1.3.3. □

A topological space $X$ is **separable** if there exists $X_0 \subset X$, countable and dense. As opposed to Proposition 1.8.5, it is not true in general that a subspace of a separable space is separable. A crude example is obtained by taking $X = [0, 1]$ with the topology $\mathcal{T} = \{T \subset X : \ 0 \in T\} \cup \{\varnothing\}$. This is a separable space because $\{0\}$ is dense: given any $t \in [0, 1]$, any open neighbourhood of $t$ contains 0, so $\overline{\{0\}} = X$. Now let $Y = [0, 1] \setminus \{0\}$. Here every $t \in Y$ is isolated, because $\{t\} = Y \cap \{0, t\}$ is open in the relative topology; so $\{t\}$ is open and $Y$ has the discrete topology $\mathcal{P}(Y)$. As $Y$ is uncountable, it does not have a countable dense subset, i.e. $Y$ is not separable. It is also possible for a Hausdorff space to be separable (even normal) and have a non-separable subspace (Exercise 1.8.10).

Given topological spaces $X$ and $Y$, we can naturally induce a topology on $X \times Y$ that proceeds from the respective topologies of $X$ and $Y$. This process is not new, as the canonical topology on $\mathbb{R}^2$ is obtained by considering $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ and constructing the product topology as we will do next. The **product topology** on $X \times Y$ is the topology generated by the sets ("rectangles") $V \times W$, where $V \subset X$ and $W \subset Y$ are open. Then

**1.8.9. Proposition.**

*Let $X, Y$ be topological spaces, and consider the product topology on $X \times Y$. Then*

*(i) the canonical projections $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ are continuous;*

*(ii) $(x_k, y_j) \to (x, y)$ if and only if $x_j \to x$, $y_j \to y$.*

If we now have arbitrarily many topological spaces $\{X_j\}$ and we want to put a topology on the Cartesian product $\prod_j X_j$, things are slightly more complicated, because what we did for two—which generalizes obviously for finitely many—can be generalized to arbitrarily many, but depending on how one does it, the results are different. Repeating the argument from above, that is constructing the topology as that generated by the Cartesian products $\prod_j V_j$ with $V_j \subset X_j$ open, gives us the **box topology**, which is not as good as it can be. Instead, the **product topology** is defined to be given by pointwise convergence, or equivalently as the coarsest topology such that all the projections $p_j$ are continuous. A huge advantage of this last choice is Tychonoff's Theorem 1.8.24 below.

**1.8.3. Continuous Functions.** Given topological spaces $X, Y$, a function $f : X \to Y$ is **continuous** if for any open $V \subset Y$ we have that $f^{-1}(V)$ is open. Continuity at a point $x_0$ can be defined by restricting the above relation to open sets $V$ that contain $f(x_0)$.

***1.8.10. Proposition.***

*Let $X, Y$ be topological spaces, $f : X \to Y$. The following statements are equivalent:*

*(i) $f$ is continuous;*

*(ii) $f$ is continuous at every point;*

*(iii) for any $A \subset X$, $f(\overline{A}) \subset \overline{f(A)}$;*

*(iv) for any closed subset $C \subset Y$, $f^{-1}(C)$ is closed;*

*(v) for every $x$, and for any net $\{x_j\}$ such that $x_j \to x$ we have $f(x_j) \to f(x)$.*

*When $X$ is a metric space, sequences can be used instead of nets.*

*Proof.* Exercise 1.8.22. $\square$

The converse of (v) in Proposition 1.8.10 does not hold in general, in the sense that it might be possible to have $f(x_j) \to f(x)$ for all continuous functions $f$ while $\{x_j\}$ does not converge to $x$. How could this be? Well, a topological space may have very few continuous functions. As an extreme example let $X = \mathbb{R}$ with the usual topology and $\mathcal{T} = \{X, \varnothing\}$. Then the only continuous functions are the constant functions, and so $f(x_j) \to f(x)$ for any net $\{x_j\}$ and any $x \in X$. For nice enough spaces (particularly, metric), though, the converse does hold; see Corollary 2.6.7 and Remark 2.6.8 for a more subtle counterexample.

If $f : X \to Y$ and $g : Y \to Z$ are continuous, then $g \circ f$ is continuous. A restriction of a continuous function is continuous.

A function $f : X \to Y$ between two topological spaces is **open** if $f(V)$ is open for all open $V \subset X$. This clearly feels related to continuity, but continuous functions need not be open maps; for instance any constant $f : X \to \mathbb{R}$ is very continuous but not open. When $f$ is bijective, $f$ is open if and only if $f^{-1}$ is continuous.

A function $f : X \to Y$ is a **homeomorphism** if $f$ is bijective, and both $f$ and $f^{-1}$ are continuous. This makes $X$ and $Y$ indistinguishable as topological spaces, since in that situation $V \subset X$ is open if and only if $f(V)$ is open. For instance, with the usual topology of the real line, $(0,1)$ and $(1,\infty)$ are homeomorphic via $f(t) = \frac{1}{t}$. Or $\mathbb{R}$ and $(0,\infty)$ are homeomorphic via the exponential. Meanwhile, $(0,1)$ and $[0,1]$ are not homeomorphic, as the latter is compact while the former isn't. For another example, $[0,1]$ is not homeomorphic to $\mathbb{T}$, as in the former we can remove a point and make it not connected, while removing a point from the latter does not affect connectedness. An injective continuous function with compact domain is a homeomorphism onto its image (Exercise 1.8.38).

A function $f : X \to \mathbb{R}$ is **upper semicontinuous** at $x_0$ if $\limsup_{x\to x_0} f(x) \leq f(x_0)$; equivalently, if for all $y > f(x_0)$ there exists an open neighbourhood $V$ of $x_0$ such that $f(z) < y$ for all $z \in V$. Similarly, $f$ is **lower semicontinuous** at $x_0$ if $\liminf_{x\to x_0} f(x) \geq f(x_0)$; equivalently, if for all $y < f(x_0)$ there exists a neighbourhood $V$ of $x_0$ such that $f(z) > y$ for all $z \in V$. The intuition is that upper semicontinuous means that if $f : \mathbb{R} \to \mathbb{R}$ has a gap discontinuity at $x_0$, then $f(x_0) \geq \lim_{x\to x_0^+} f(x)$; and $f$ is lower semicontinuous at $x_0$ if $f(x_0) \leq \lim_{x\to x_0^-} f(x)$.

### *1.8.11. Proposition.*

*Let $X$ be a topological space and $f : X \to \mathbb{R}$. Then*

*(i) $f$ is upper semicontinuous at $x_0$ if and only if $f^{-1}[f(x_0), \infty)$ is closed;*

*(ii) $f$ is lower semicontinuous at $x_0$ if and only if $f^{-1}(-\infty, f(x_0)]$ is closed;*

*(iii) $f$ is both upper and lower semicontinuous at $x_0$ if and only if $f$ is continuous at $x_0$.*

For global semicontinuity we have

### *1.8.12. Proposition.*

*Let $X$ be a topological space and $f : X \to \mathbb{R}$. Then*

*(i) $f$ is upper semicontinuous everywhere if and only if $f^{-1}[y, \infty)$ is closed for all $y \in \mathbb{R}$;*

(ii) *$f$ is lower semicontinuous everywhere if and only if $f^{-1}(-\infty, y]$ is closed for all $y \in \mathbb{R}$;*

(iii) *$f$ is both upper and lower semicontinuous if and only if $f$ is continuous.*

One consequence of this is

***1.8.13. Corollary.***

*Let $X$ be a topological space and $\mathcal{R} \subset C(X)$. Then the function $h$ defined by*

$$h(t) = \sup\{f(t) : \ f \in \mathcal{R}\}$$

*is lower semicontinuous. Similarly, the infimum is upper semicontinuous.*

*Proof.* Fix $y \in \mathbb{R}$ and consider the set $h^{-1}(-\infty, y]$. Suppose that there is a net $\{s_j\} \subset X$ such that $s = \lim_j s_j$ where $s_j \in h^{-1}(-\infty, y]$ for all $j$. Fix $\varepsilon > 0$. By definition of $h$, there exists $f \in \mathcal{R}$ such that $h(s) \le f(s) + \varepsilon$. And by the continuity of $f$ there exists $j$ such that $f(s) \le f(s_j) + \varepsilon$. Then

$$h(s) \le f(s) + \varepsilon \le f(s_j) + 2\varepsilon \le y + 2\varepsilon.$$

As $\varepsilon$ was arbitrary, it follows that $h(s) \le y$ and so $s \in h^{-1}(-\infty, y]$. So $h^{-1}(-\infty, y]$ is closed and $h$ is lower semicontinuous by Proposition 1.8.12. The assertion for the infimum follows from the elementary fact that $h$ is lower semicontinuous if and only if $-h$ is upper semicontinuous. $\square$

Here is a result we will need a long time from now.

***1.8.14. Lemma.***

*Let $T$ be a Hausdorff topological space and $f : T \to T$ continuous and surjective. If $f \neq \mathrm{id}_T$, then there exists $E \subsetneq T$, closed, with $T = E \cup f(E)$.*

*Proof.* Suppose that there exists $t_0 \in T$ with $f(t_0) \neq t_0$. Then we can choose disjoint open neighbourhoods $V, W$, with $t_0 \in V$ and $f(t_0) \in W$. Let $E = T \setminus \big(V \cap f^{-1}(W)\big)$. As $t_0 \in V \cap f^{-1}(W)$, $t_0 \notin E$. Given $t \in T \setminus E = V \cap f^{-1}(W)$, the surjectivity allows us to write $t = f(s)$ for some $s \in T$. If $s \in V \cap f^{-1}(W)$ then $t = f(s) \in W$ and $t \in V$, a contradiction. That is, $s \notin T \setminus E$, so $s \in E$ and $t \in f(E)$. We have shown that $t \in E \cup f(E)$. $\square$

**1.8.4. Bases and Subbases.** In most topological arguments in metric spaces we do not use arbitrary open sets, but rather use balls, since every open set is a union of balls, and for any point in an open set there is a ball, inside the set, containing the point. In topological spaces we will also often have the need to deal with a family of simpler open sets that generate the topology. A **base** for $X$ is a family $\mathcal{B}$ of open sets such that

(1) for each $V \subset X$ open, $\displaystyle\bigcup_{B\in\mathcal{B}, B\subset V} B = V$;

(2) given $B_1, B_2 \in \mathcal{B}$, for each $x \in B_1 \cap B_2$ there exists $B_3 \in \mathcal{B}$ with $x \in B_3 \subset B_1 \cap B_2$.

Given $x \in X$, a **local base** is a family $\mathcal{B}_x$ of open sets such that for any open set $V$ with $x \in V$, there exists $B \in \mathcal{B}_x$ with $x \in B \subset V$. When $X$ has a countable local base at each point, it is said to be **first countable**. When $X$ has a countable base, we say it is **second countable**.

The computations necessary to solve Exercise 1.8.1 are what is needed to show that the balls in a metric space form a base for the topology.

It is also common to use a **subbase** of the topology $\mathcal{T}$. This is a family $\mathcal{S}$ of subsets of $\mathcal{T}$ that generates $\mathcal{T}$ (that is, $\mathcal{T}$ is the smallest topology that contains $\mathcal{S}$). Equivalently, $\mathcal{S} \subset \mathcal{T}$ is a subbase if the family of finite intersections of sets in $\mathcal{S}$ is a base for $\mathcal{T}$. A typical example of a subbase would be, on the real line, the family of subsets of the form $(a, \infty)$ or $(-\infty, b)$.

For a subset $X_0$ of the topological space $X$, the **subspace topology** (also, **relative topology**) is the topology $\{V \cap X_0 : \ V \in \mathcal{T}\}$.

**1.8.5. Connectedness.** A topological space $X$ is **connected** if $X = V_1 \cup V_2$ with $V_1, V_2$ open and disjoint, implies that either $V_1$ or $V_2$ is empty. A subset of $X$ is connected if it is connected in its subspace topology. It is not hard to show that for any topological space $X$, every $x \in X$ belongs to a unique maximally connected subset. Indeed, we use Zorn's Lemma on the family of connected subsets of $X$ that contain $x$; if $\{E_j\}$ is a chain of connected subsets such that $x \in E_j$ for all $j$, then $E = \bigcup_j E_j$ is connected. For if $E = V_1 \cup V_2$ with $V_1, V_2$ disjoint and open in the subspace topology of $E$, then there exist open sets $W_1, W_2$ with $V_1 = E \cap W_1$, $V_2 = E \cap W_2$. For any $j$ we have $E_j = (E_j \cap V_1) \cup (E_j \cap V_2) = (E_j \cap W_1) \cup (E_j \cap W_2)$ contradicting that $E_j$ is connected. So $E$ is an upper bound for $\{E_j\}$. By Zorn's Lemma, there exists $E_x \subset X$, maximal connected, with $x \in E_x$. The maximality guarantees that $E_x$ is unique: if we define a relation by $x \sim y$ if $y \in E_x$, this is an equivalence relation (Exercise 1.8.35). This means that each element in $X$ belongs to a unique **connected component**, and $X$ is the disjoint union of its connected components. As the closure of a connected set is connected (Exercise 1.8.33) and connected components are maximal, it follows that each connected component is closed. In the case where there are finitely many connected components, each is also open, as its complement is a finite union of closed which is closed. Sets which are both open and closed are called **clopen**.

Here is a result that will be needed many, many pages from now.

***1.8.15. Proposition.***

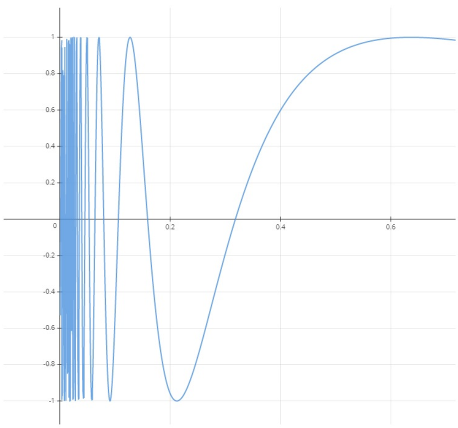
*Let $(X, d)$ and $(Y, d')$ be metric spaces and $f : X \to Y$ continuous. Show that if $f$ is **locally constant** (that is, for each $x \in X$ there exists $r > 0$ such that $f$ is constant on $B_r(x)$) then $f$ is constant on each connected component of $X$.*

*Proof.* Assume without loss of generality that $X$ is connected. Fix $x_0 \in X$ and let $V = \{x \in X : f(x) = f(x_0)\}$. If $x \in V$, by hypothesis there exists $r > 0$ with $B_r(x) \subset V$; so $V$ is open. We also have $V = f^{-1}(\{f(x_0)\})$, a preimage of a closed set through a continuous function; so $V$ is closed. Then $V = X$ for otherwise $X = V \cup V^c$ would contradict the connectedness of $X$. $\square$

A topological space is **path connected** if given any two points $x, y \in X$ there exists a continuous function $f : [0, 1] \to X$ with $f(0) = x$, $f(1) = y$. Every path-connected space is connected, but not every connected space is path-connected. The canonical example of a connected but not path-connected space is

$$X = \{(0, y) : \ y \in [-1, 1]\} \cup \left\{ \left(x, \sin \tfrac{1}{x}\right) : \ x \in [0, 1] \right\} \subset \mathbb{R}^2,$$

pictured below.




**1.8.6. Compactness.** A subset $K \subset X$ is **compact** if any open cover of $K$ admits a finite subcover. That is, if $K \subset \bigcup_j V_j$ with $V_j$ open for all $j$, then there exist $j_1, \ldots, j_n$ with $K \subset \bigcup_{k=1}^n V_{j_k}$.

**1.8.16. Lemma.**

*A closed subset of a compact set is compact. In a Hausdorff space $X$, every compact subset of $X$ is closed. And, in a metric space, a compact set has to be bounded.*

*Proof.* Let $K \subset X$ be compact, and $C \subset K$ be closed. Let $\mathcal{V}$ be an open cover for $C$. Then
$$K \subset (X \setminus C) \cup \bigcup_{V \in \mathcal{V}} V.$$
Since $K$ is compact, there exist $V_1, \ldots, V_n \in \mathcal{V}$ with $K \subset (X \setminus C) \cup V_1, \cdots \cup V_n$. So $C \subset V_1 \cup \cdots \cup V_n$ and $C$ is compact. Now assume that $X$ is Hausdorff and $K \subset X$ is compact. Fix $x \in X \setminus K$; for each $z \in K$, there exist disjoint open sets $V_z$ and $W_z$ such that $x \in V_z$, $z \in W_z$. Since $K \subset \bigcup_{z \in K} W_z$, there exist $z_1, \ldots, z_m \in K$ such that $K \subset W = \bigcup_{k=1}^m W_{z_k}$. The set $W$ is open and $x \notin W$. And $x \in V = \bigcap_{k=1}^m V_{z_k}$ which is open, and $V \cap W = \varnothing$, so $V \subset X \setminus K$. We have shown that $x$ is an interior point of $X \setminus K$, and so $X \setminus K$ is open, which means that $K$ is closed.

If $K$ is not bounded, then the family of open balls $\{B_n(0)\}_{n \in \mathbb{N}}$ provides an open cover without finite subcover. □

We will need this lemma, which shows a typical way in which open covers are used; though it is more common in applications to use other equivalent characterizations of compactness. It is a particular case of Tychonoff's Theorem, but we offer an ad-hoc proof here.

**1.8.17. Lemma.**

*Let $X, Y$ be topological spaces and consider the product topology on $X \times Y$. If $K \subset X$, $H \subset Y$ are compact, then $K \times H$ is compact.*

*Proof.* Let $\mathcal{V}$ be an open cover of $K \times H$. For each $(x, y) \in K \times H$ there exists $V_{x,y} \in \mathcal{V}$ with $(x, y) \in V_{x,y}$. By definition of the product topology, there exist open sets $U_{x,y} \subset X$, $W_{x,y} \subset Y$ such that $(x, y) \in U_{x,y} \times W_{x,y} \subset V_{x,y}$. For fixed $x \in K$, the family $\{W_{x,y}\}_{y \in H}$ is an open cover of $H$. So there exist $y_1(x), \ldots, y_{m(x)}(x) \in H$ with $H \subset W_{x,y_1(x)} \cup \cdots \cup W_{x,y_{m(x)}(x)}$. Let
$$U_x = \bigcap_{j=1}^{m(x)} U_{x,y_j(x)}.$$

Then $U_x$ is open, and $K \subset \bigcup_{x \in K} U_x$. By compactness of $K$, there exist $x_1, \dots, x_n \in K$ such that $K \subset \bigcup_{k=1}^n U_{x_k}$. Then

$$K \times H \subset \bigcup_{k=1}^{n} \bigcup_{j=1}^{m(x_k)} U_{x_k} \times W_{x_k, y_j(x_k)} \subset \bigcup_{k=1}^{n} \bigcup_{j=1}^{m(x_k)} V_{x_k, y_j(x_k)}$$

and we have a finite subcover. □

In a metric space, compactness of $K$ is equivalent to the existence, for any $\varepsilon > 0$, of an $\varepsilon$-net: that is, finitely many points $x_1, \dots, x_m$ such that $K \subset \bigcup_{j=1}^m B_\varepsilon(x_j)$, see Lemma 1.8.25. In the space $\mathbb{R}^n$, the usual stomping ground where one learns analysis, compactness is characterized by a well-known result:

***1.8.18. Theorem. (Heine–Borel)***

*Let $K \subset \mathbb{R}^n$. Then $K$ is compact if and only if $K$ is closed and bounded.*

*Proof.* If $K$ is compact, then it is closed and bounded by Lemma 1.8.16..

Suppose that $K$ is closed and bounded. Then there exists a closed interval $B = [-c, c]^n$ such that $K \subset B$. If we show that $B$ is compact, then $K \subset B$ is compact by Lemma 1.8.16, being a closed subset of compact. So it is enough to show that $B$ is compact; as compactness is preserved by homeomorphisms, it is actually enough to show that $[0,1]^n$ is compact. In fact, by Lemma 1.8.17, it is enough to show that $[0,1]$ is compact. For this, a usual argument is as follows. Let $\mathcal{V}$ be an open cover of $[0,1]$. Let

$$T = \{t \in [0,1] : \quad \text{a finite subcollection of } \mathcal{V} \text{ covers } [0,t]\}.$$

Then $T \neq \varnothing$ since $0 \in T$. Let $s = \sup T$. Let $V_0 \in \mathcal{V}$ with $s \in V_0$. Since $V_0$ is open and $s \in V_0$, there exists $\delta > 0$ such that $(s - \delta, s + \delta) \subset V_0$. And by definition of supremum there exists $s' \in T$ with $s' \in (s - \delta, s)$. As $s' \in T$, the interval $[0, s']$ admits a finite subcover $V_1, \dots, V_n$ of $\mathcal{V}$. Then $V_0, V_1, \dots, V_n$ is a finite subcover of $[0, s]$, showing that $s \in T$. The open sets $V_0, \dots, V_n$ cover $[0, s + \delta)$, so $[0, s + \delta/2] \subset V_0 \cup \dots \cup V_n$; if $s < 1$ this would be a contradiction, as $s = \max T$. Thus $s = 1$. □

Since topologically $\mathbb{C}^n = \mathbb{R}^{2n}$, we can replace $\mathbb{R}$ by $\mathbb{C}$ in Theorem 1.8.18.

Because of Heine–Borel, and depending on the path taken to learn analysis, students may have the ingrained knowledge—and this is how it is often used—that a closed and bounded set is compact. But this is not true in general; see for instance Theorems 4.4.8 and 5.2.9.

A family $\{X_j\}_{j \in J}$ of sets has the **finite intersection property** if for any $j_1, \dots, j_n \in J$ we have $\bigcap_{k=1}^n X_{j_k} \neq \varnothing$.

***1.8.19. Proposition.***

*Let $X$ be a topological space. The following statements are equivalent:*

*(i)* *$X$ is compact;*

*(ii)* *any net in $X$ admits a convergent subnet;*

*(iii)* *if $K_j \subset K$ is closed for all $j$ and $\{K_j\}$ has the finite intersection property, then $\bigcap_j K_j \neq \varnothing$.*

*Proof.* (i) $\implies$ (iii) Suppose that there exists a family $\{K_j\}$ of closed sets such that such that any finite intersection is non-empty but $\bigcap_j K_j = \varnothing$. Then

$$X = \varnothing^c = \Big(\bigcap_j K_j\Big)^c = \bigcup_j K_j^c$$

so $\{K_j^c\}$ is an open cover for $X$; this means that there exist $j_1, \ldots, j_m$ with $X \subset \bigcup_{k=1}^m K_{j_k}^c$. Taking complements we get $\bigcap_{k=1}^m K_{j_k} = \varnothing$, a contradiction.

(iii) $\implies$ (ii) Let $\{x_j\}_{j \in J}$ be a net in $X$. Define

$$R_j = \overline{\{x_k :\ k \geq j\}}.$$

Then $\{R_j\}$ is a net of closed sets that has the finite intersection property:

$$R_{j_1} \cap \cdots \cap R_{j_m} \supset \{x_j :\ j \geq \max\{j_1, \ldots, j_m\}\}.$$

So $\bigcap_j R_j \neq \varnothing$. Let $x \in \bigcap_j R_j$. For any open set $V$ with $x \in V$ we have $x \in \bigcap_j (R_j \cap V)$, so for every $j$ we have $x \in R_j \cap V$ and thus there exists $k_{V,j} \in J$ such that $k_{V,j} \geq j$ and $x_{k_{V,j}} \in V$. On the set of triples $(V, k_{V,j}, j)$, indexed by all open subsets of $X$ and $J$, we consider the partial order $(V, k_{V,j}, j) \leq (W, k_{W,h}, h)$ if $V \supset W$, $k_{V,j} \leq k_{W,h}$ and $j \leq h$. This way we obtain a subnet $\{x_{k_{V,j}}\}$ of $\{x_j\}$ that converges to $x$: given $V$ open with $x \in V$ and $j \in J$, for any $(W, k_{W,h}, h) \geq (V, k_{V,j}, j)$ we have that $x_{k_{W,h}} \in V$.

(ii) $\implies$ (i) Let $\{V_j\}_{j \in J}$ be an open cover of $X$ with no finite subcover. Define $\mathcal{J} = \{J_0 \subset J :\ |J_0| < \infty\}$. As there is no finite subcover, for each $H \in \mathcal{J}$ there exists $x_H \in X$ such that $x_H \notin V_j$ for all $j \in H$. The set $\mathcal{J}$ is directed when ordered by inclusion, so we may consider the net $\{x_H\}_{H \in \mathcal{J}}$. Given $x \in X$ there exists $j \in J$ such that $x \in V_j$. For all $H$ with $j \in H$ we have that $x_H \notin V_j$; so the net is separated from $x$ and so it cannot have a subnet convergent to $x$; as we can do this for any $x$, the net does not admit a convergent subset. $\square$

*1.8.20. Remark*. Proposition 1.8.19 allows for a different notion of compactness, which is **sequential compactness**. This means relaxing the definition to assert that every *sequence* has a convergent subsequence. This notion does not agree with compactness in general, though it does in first countable spaces; in particular, in metric spaces.

***1.8.21. Proposition.***

*Let $X$ be a compact metric space. Then $X$ is complete.*

*Proof.* Let $\{x_n\} \subset X$ be a Cauchy sequence. Since $X$ is compact, by Proposition 1.8.19 there is a convergent subset. Because we are in a metric space we can obtain a convergent subsequence out of the convergent subset. So there exists $x \in X$ and $\{x_{n_k}\}$ such that $x_{n_k} \to x$. But then

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x).$$

By choosing $n$ and $k$ big enough, given $\varepsilon > 0$ we can get $d(x_n, x) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, and so $x_n \to x$. $\square$

Here is an application that we promised on Remark 1.6.24.

***1.8.22. Corollary.*** *([Mun00], Theorem 27.7)*

*Let $X$ be a nonempty compact Hausdorff space with no isolated points. Then $X$ is uncountable.*

*Proof.* Let $g : \mathbb{N} \to X$ be any function.

**Claim.** For any $x \in X$ and $U \subset X$ open, there exists $V \subset U$, open and nonempty, with $x \notin \overline{V}$.

Assuming the claim, there exists $V_1 \subset X$ open with $g(1) \notin \overline{V_1}$. Inductively, given $V_{n-1}$ open and nonempty, there exists $V_n \subset V_{n-1}$ open and nonempty with $g(n) \notin \overline{V_n}$. Now the sequence $\overline{V_1} \supset \overline{V_2} \supset \cdots$ of compact sets has the finite intersection property, being monotone and made out of nonempty sets. By Proposition 1.8.19, there exists $x \in \bigcap_n V_n$. By construction, $x \neq g(n)$ for all $n$. So $g$ is not surjective. In particular, $|X| > |\mathbb{N}|$.

It remains to prove the claim. Let $y \in U$ with $y \neq x$. The existence of $y$ follows from $U \neq \varnothing$ if $x \notin U$; and if $x \in U$, it follows from the fact that $x$ is not isolated, so any neighbourhood of $x$ has to contain at least one other point. Because $X$ is Hausdorff, there exist open sets $W_1, W_2 \subset X$, disjoint, with $x \in W_1$ and $y \in W_2$. Let $V = U \cap W_2$. Then $V$ is open, nonempty since $y \in V$, and $x \notin \overline{V}$ since $x \in W_1$ and $W_1 \cap W_2 = \varnothing$. $\square$

One can use Corollary 1.8.22 to show that $|[0,1]| > |\mathbb{N}|$, and then $|\mathbb{R}| \geq |[0,1]| > |\mathbb{N}|$.

*1.8.23. Remark.* It should be noted that the idea in Corollary 1.8.22 is not radically different from the diagonal argument from Theorem 1.6.22. Rather, it is a smart generalization. The goal of the diagonal argument is to construct a point $x$ with $x \neq g(n)$ for all $n$. That's precisely the $x \in \bigcap_n V_n$ in Corollary 1.8.22 . If $g(1) = 0.123\cdots$, when you prescribe the first decimal digit of $x$ to be 5, say, you are prescribing $V_1 = (0.4, 0.6)$. Then if you prescribe the second digit to be 4, now $V_2 = (0.43, 0.57)$. Continuing this way you get the sequence of open sets $\{V_n\}$ with $x_n \notin V_n$ and $x \in \bigcap_n V_n$. We will see on Remark 2.3.13 an argument that does not rely on the diagonal at all.

An important fact about compactness is the following:

***1.8.24. Theorem. (Tychonoff)***

*Let $\{X_j\}$ be a family of compact topological spaces. On the Cartesian product $\prod_j X_j$ consider the pointwise topology, i.e. $f_\alpha \to f$ if $f_\alpha(j) \to f(j)$ for all $j$. Then $\prod_j X_j$ is compact.*

When our topological space is metric, another characterization of compactness is possible. The centres of the balls in Lemma 1.8.25 are often called an **$\varepsilon$-net**.

***1.8.25. Lemma.***

*In a complete metric space, a closed subset $K$ is compact if and only if for every $\varepsilon > 0$ there exists a finite number of balls of radius $\varepsilon$ that covers $K$.*

*Proof.* If $K$ is compact and $\varepsilon > 0$, we have

$$K \subset \bigcup_{k\in K} B_\varepsilon(k).$$

By compactness, there exist $k_1, \ldots, k_r \in K$ with

$$L \subset \bigcup_{j=1}^{r} B_\varepsilon(k_j).$$

Conversely, suppose that for every $\varepsilon > 0$ there exists a finite number of balls of radius $\varepsilon$ that covers $K$. Let $\{x_j\} \subset K$ be a net. By hypothesis we can cover $K$ with finitely many balls of radius 1. At least one of them will contain a cofinal subnet (otherwise, there would be a maximum index); pick such ball, call it $B_1$. Choose $j_1$ in the subnet such that $x_{j_1} \in B_1$. We can cover $K$ with finitely many balls of radius $1/2$; let $B_2$ be one such ball will contain a cofinal subnet of $\{x_j : j \geq j_1\}$; choose one such $x_{j_2}$. Inductively, we end up with a sequence $\{x_{j_n}\}$ (with $j_{n+1} \geq j_n$, so it is a subnet of $\{x_j\}$) such that if $n \geq n_0$ then $x_{j_n} \in B_{1/n_0}$. This implies that the sequence is Cauchy; as the space is complete and $K$ is closed, thus complete, the sequence is convergent. By Proposition 1.8.19, $K$ is compact. □

Lemma 1.8.25 can fail without completeness (Exercise 1.8.12).

If $f : X \to Y$ is continuous and $X$ is compact, then $f(X)$ is compact (Exercise 1.8.37). As compact subsets of a metric space are closed and bounded, this shows that a continuous real-valued function on a topological space will achieve its maximum and its minimum on any compact set.

A topological space $X$ is **locally compact** if every point admits an open neighbourhood with compact closure. Local compactness is key for the existence of sufficiently many good continuous functions, as we will see in Section 2.6.

Since compactness is so nice, it is often a good idea to embed our non-compact space into a compact one; similar to how we learned above how to embed any metric space into a complete one.

**1.8.26. Definition.**

*Let $T$ be a Hausdorff topological space. A compact Hausdorff space $\tilde{T}$ is a **compactification** for $T$ if there exists $\gamma : T \to \tilde{T}$, homeomorphic onto its image, and such that $\gamma(T)$ is dense in $\tilde{T}$.*

The minimal way to obtain a compactification is the **one-point compactification** (also known as **Alexandroff's compactification**, see below). Much later we will see another kind of compactification of a locally compact Hausdorff space, maximal in a sense: the Stone–Čech compactification (Theorem 7.4.14).

The definition of compactification above precludes compactifying an already compact Hausdorff space with anything other than itself. For if $T$ is compact Hausdorff and $\gamma : T \to \tilde{T}$ is a homeomorphism onto its image with $\tilde{T}$ Hausdorff, then $\gamma(T)$ is a compact subset of $\tilde{T}$, hence closed. Thus if $\tilde{T} \setminus \gamma(T) \neq \varnothing$ we have that $\gamma(T)$ is not dense.

The symbol $\infty$ below is just that, just a symbol, and we could have used any other symbol. It is common to use $\infty$, though, inspired by the case where $T = [0, \infty)$. When $T = \mathbb{R}$, the one-point compactification identifies $-\infty$ and $+\infty$; in fact, one can show that $\mathbb{R}_\infty$ is homeomorphic with $\mathbb{T}$ (Exercise 1.8.44). It should be clear that other compactifications are possible. For instance, it is not hard to compactify $\mathbb{R}$ as $\mathbb{R} \cup \{-\infty, \infty\}$ (Exercise 1.8.41).

***1.8.27. Proposition. (Alexandroff's Compactification)***

*Let $T$ be a non-compact locally compact Hausdorff space. Then there exists a topology on $T_\infty = T \cup \{\infty\}$ such that $T_\infty$ is compact Haussdorff, and the subspace topology on $T$ agrees with the original topology of $T$.*

*Proof.* On $T_\infty = T \cup \{\infty\}$, we consider the following topology:

$$\mathrm{Top}\Big\{\{V \subset T, \text{ open}\} \cup \{(T \setminus K) \cup \{\infty\} : \; K \subset T \text{ compact}\}\Big\}.$$

As the sets $T \setminus K$ are open, the subspace topology is clearly the topology of $T$. So it remains to show that $T_\infty$ is compact. Let $\{T_j\}_{j\in J}$ be an open cover of $T_\infty$. At least one of these open sets contains $\infty$, so it is of the form $(T \setminus K) \cup \{\infty\}$ for some compact $K$. If we remove this set from the list and we remove $\infty$ from all other sets in the cover, we get a cover $\{T'_j\}$ such that $T_{j_0} = (T \setminus K) \cup \{\infty\}$ and $\infty \notin T'_j$ for all $j \neq j_0$. As $\mathbb{R} \subset \bigcup_j T'_j$, we have that $K \subset \bigcup_{j\neq j_0} T'_j$. Because $K$ is compact, there exist $j_1, \dots, j_m$ such that $K \subset T_{j_1}, \dots, T_{j_m}$. Then $T_\infty = T_{j_1} \cup \cdots \cup T_{j_m} \cup ((T \setminus K) \cup \{\infty\})$, a finite subcover, showing that $T_\infty$ is compact. ☐

A naive question to ask is what happens if one remove the local compactness requirement. What happens is that closed subsets of compact subsets are compact, so if $T$ is not locally compact there exists $t \in T$ such that $\overline{V}$ is not compact for any open $V$ with $t \in V$. If we now think of $T$ inside a compactification, we can choose $V$ that separates $t$ from $\infty$, and so $\overline{V}$ is a closed subset of a compact space, hence compact, a contradiction.

### 1.8.7. Exercises.

(1.8.1) Let $X$ be a metric space, $x \in X$ and $r > 0$. Show that $B_r(x)$ is open.

(1.8.2) Let $X$ be a metric space. Let $\{A_j\}$ be a collection of open sets, and let $\{B_k\}$ be a collection of closed sets. Show that $\bigcup_j A_j$ is open, and that $\bigcap_k B_k$ is closed.

(1.8.3) Find an example of a metric space $X$ and a collection $\{A_j\}$ of open sets such that $\bigcap_j A_j$ is not open. Find also an example of a collection $\{B_k\}$ of closed sets such that $\bigcup_k B_k$ is not closed.

(1.8.4) Let $X$ be a topological space and $V, W \subset X$ open and disjoint. Show that $\overline{V} \cap W = \varnothing$.

(1.8.5) Show that a metric space is normal.

(1.8.6) Let $X$ be a topological space. Show that the following statements are equivalent:

(i) $X$ is normal;

(ii) given $K \subset V$ with $K$ closed and $V$ open, there exists $W \subset X$, open, with $K \subset W \subset \overline{W} \subset V$.

(1.8.7) Prove Proposition 1.8.6.

(1.8.8) Let $M$ be a separable metric space and $X \subset M$ be uncountable. Show that $X$ has infinitely many accumulation points.

(1.8.9) Let $X = \mathbb{R}$ and $\mathcal{T} = \operatorname{Top}\{[a,b) : \ a,b \in \mathbb{R}, \ a < b\}$. The topological space $(\mathbb{R}, \mathcal{T})$ is called the **Sorgenfrey Line**, and $\mathcal{T}$ is called the **lower limit topology**.

(i) Show that $\mathcal{T}$ is finer than the usual topology on $\mathbb{R}$.

(ii) Show that $[a,b)$ is both open and closed for all $a < b$.

(iii) Show that $(\mathbb{R}, \mathcal{T})$ is normal.

(iv) Show that $x_n \to x$ if and only if there exists $n_0$ such that $x_n \geq x$ for all $n \geq n_0$, and $x_n \to x$ in the usual topology.

(v) Show that if $K \subset \mathbb{R}$ is compact then $K$ is countable.

(1.8.10) Let $(\mathbb{R}, \mathcal{T})$ be the Sorgenfrey Line and consider the topological space $(\mathbb{R}^2, \mathcal{T} \times \mathcal{T})$. This is called the **Sorgenfrey Plane**.

(i) Show that the Sorgenfrey Plane is separable.

(ii) Consider the set $Y = \{(x, -x) : \ x \in \mathbb{R}\} \subset \mathbb{R}^2$. Show that $Y$ is not separable.

(iii) Conclude that the Sorgenfrey Plane is not metric.

(iv) Show that $Y_0 = \{(x, -x) : \ x \in \mathbb{Q}\}$ and $Y \setminus Y_0$ are closed, and use this information to show that $(\mathbb{R}, \mathcal{T})$ is not normal.

(1.8.11) Show that a discrete compact topological space is finite.

(1.8.12) Show that completeness is actually required in Lemma 1.8.25.

(1.8.13) Using the $\varepsilon$-$\delta$ definition of continuity in a metric space, show that everywhere continuity of $f : X \to Y$ is equivalent to saying that $f^{-1}(E)$ is open in $X$ for every open set $E \subset Y$.

(1.8.14) Show that $f : X \to Y$ is continuous if and only if $f^{-1}(E)$ is open for every $E$ in a subbase for $X$.

(1.8.15) Let $X$ be a topological space and $H \subset X$ a subset. Show that $1_H$ is continuous if and only if $H$ is clopen.

(1.8.16) Let $(X, d)$ be a metric space and $\{f_n\}$ a sequence of continuous functions such that $f_n \to f$ uniformly. Show that $f$ is continuous.

(1.8.17) Let $(M, d)$ and $(N, d')$ be metric spaces and $f : M \to N$ a function. As mentioned, $f$ is continuous at $x \in M$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $d(y, x) < \delta \implies d'(f(y), f(x)) < \varepsilon$. When $\delta$ does not depend on $\mathcal{X}$, we say that $f$ is **uniformly continuous**. Show that if $M$ is compact then $f$ is uniformly continuous.

(1.8.18) Let $X = \{1,2,3\}$ with the topology $\{\varnothing, X, \{1\}, \{2,3\}\}$. Show that $f : X \to \mathbb{R}$ is continuous if and only if $f(2) = f(3)$.

(1.8.19) Let $a, b \in \mathbb{R}$ and $f : (a,b) \to \mathbb{R}$ be continuously differentiable. Show that $f$ is piecewise monotonic, i.e., there exists disjoint intervals $(a_k, b_k)$ such that $[a,b] = \bigcup_k [a_k, b_k]$, $f$ is monotone on each $[a_k, b_k]$, and there exist intervals $(a_k', b_k') \subset (a_k, b_k)$ such that $f$ is strictly monotone on $(a_k', b_k')$ and constant on $(a_k, a_k')$ and on $(b_k', b_k)$.

(1.8.20) Let $X = \{1,2,3\}$. Show that $\mathscr{T} = \{\varnothing, X, \{1,2\}, \{2\}, \{2,3\}\}$ is a topology, and that it is not Hausdorff.

(1.8.21) Let $X = \{1,2,3\}$ with the topology $\{\varnothing, X, \{1,2\}, \{2\}, \{2,3\}\}$. Show that $f : X \to \mathbb{R}$ is continuous if and only if $f$ is constant.

(1.8.22) Prove Proposition 1.8.10.

(1.8.23) Show that an interval $(a,b) \subset \mathbb{R}$ is connected.

(1.8.24) Let $(X,d)$ be a metric space. Show the **reverse triangle inequality**

$$|d(x,y) - d(y,w)| \leq d(x,w), \qquad x, y, w \in X.$$

(1.8.25) Let $X$ be a complete topological space and $C \subset X$ a closed subset. Show that $C$ is complete.

(1.8.26) Let $(X,d)$ be a metric space. Construct a completion for $X$ in the following way. Let $\tilde{X}$ be the set of Cauchy sequences in $X$, and $R$ the equivalence relation

$$(x_n)\, R\, (y_n) \iff d(x_n, y_n) \to 0.$$

On $\bar{X} = \tilde{X}/R$ define

$$d'(\,(x_n), (y_n)\,) = \lim_n d(x_n, y_n).$$

(i) Show that $d'$ is well-defined.

(ii) Show that the map $\rho : X \to \bar{X}$ that maps $x$ to the constant sequence $(x)$ is isometric.

(iii) Show that $\rho(X)$ is dense in $\bar{X}$.

(iv) Show that $\bar{X}$ is complete.

(1.8.27) Let $(X,d)$ be a metric space and $\{x_n\}$ a Cauchy sequence. Show that $\{x_n\}$ is bounded; that is, there exists $x \in X$ and a ball $B$ centered at $x$ such that $\{x_n\} \subset B$.

(1.8.28) Let $X, Y$ be complete metric spaces with dense subsets $X_0, Y_0$ respectively. Let $\gamma : X_0 \to Y_0$ be an isometric surjection. Show that there exists a unique $\tilde{\gamma} : X \to Y$, bijective and isometric.

(1.8.29) Let $X = \mathbb{Q}$ and define

$$d(x,y) = |\arctan x - \arctan y|.$$

Show that $d$ is a distance, and find the completion of $(X,d)$.

(1.8.30) Let $X = \mathbb{R} \setminus \mathbb{Q}$, with

$$d(x,y) = \begin{cases} |x-1| + |y-1|, & x \neq y \\ 0, & x = y \end{cases}$$

Show that $d$ is a metric, and find the completion of $(X, d)$.

(1.8.31) Let $X$ be a complete metric space, and $E_1 \supset E_2 \supset \cdots$ a decreasing sequence of closed sets, such that $\lim_n \operatorname{diam}(E_n) = 0$. Show that $\bigcap_n E_n$ is nonempty and it consists of a single point. Can the "closed" condition be removed?

(1.8.32) Let $X$ be a topological space, and let $V, W \subset X$ be disjoint open subsets. Show that $\overline{U} \cap W = \varnothing$.

(1.8.33) Let $X$ be a topological space, $E \subset X$. Show that if $E$ is connected, then $\overline{E}$ is also connected.

(1.8.34) Let $X$ be a topological space and $E, F \subset X$ connected. Show that if $E \cap F \neq \varnothing$, then $E \cup F$ is connected.

(1.8.35) Let $X$ be a topological space. For each $x$, denote by $E_x$ a maximal connected subset with $x \in E_x$. Define a relation by $x \sim y$ if $y \in E_x$. Show that $\sim$ is an equivalence relation.

(1.8.36) Show that a path-connected space is connected.

(1.8.37) Let $X, Y$ be topological spaces, $f : X \to Y$ continuous, and $K \subset X$ compact. Show that $f(K)$ is compact.

(1.8.38) Let $X$ be compact Hausdorff, $Y$ a Hausdorff topological space, and $\psi : X \to Y$ continuous. Show that if $\psi$ is injective, then $\psi$ is a homeomorphism onto $\psi(X)$.

(1.8.39) Use Exercise 1.8.38 to show that if $X$ is compact Hausdorff, any weaker topology on $X$ is not Hausdorff, and any stronger topology on $X$ is Hausdorff but not compact.

(1.8.40) Let $X, Y$ be topological spaces and $f : X \to Y$ continuous. If $X_0 \subset X$ is such that $\overline{X_0}$ is compact, show that $f(\overline{X_0}) = \overline{f(X_0)}$.

(1.8.41) Show that the set $\mathbb{R} \cup \{-\infty, \infty\}$ can be given a topology such that it is a compactification of $\mathbb{R}$.

(1.8.42) Let $S, T$ be homeomorphic topological spaces such that both are locally compact Hausdorff . Show that $S_\infty$ is homeomorphic to $T_\infty$.

(1.8.43) Let $T$ be a locally compact Hausdorff space. Let $R$ and $S$ be one-point compactifications of $T$, that is $R = T \cup \{\infty_R\}$ is a compact Hausdorff space such that the restriction to $T$ agrees with the topology of $T$, and similarly for $S$. Show that $R$ and $S$ are homeomorphic; that is, the one-point compactification is unique.

(1.8.44) Show that the one-point compactification $\mathbb{R}_\infty$ of $\mathbb{R}$ is homeomorphic to the unit circle $\mathbb{T}$.

(1.8.45) Let $S, T$ be topological spaces and $f : S \to T$ be continuous and surjective and such that $f$ is not surjective when restricted to any proper closed subset of $S$. Let $U \subset S$ be open. Show that $f(U) \subset \overline{T \setminus f(S \setminus U)}$.

(1.8.46) Let $X$ be a separable metric space, and $V \subset X$ open. Show that $V$ is a countable union of balls.

CHAPTER

# 2

# Measure and Integration

Measure and Integration grew up out of attempts to address shortcomings in the Riemann integral. It ended up not only solving many of said shortcomings, but actually becoming an area on its own, and extending the notion of integral to realms that would have been unthinkable for the original Riemann integral. Besides its own worth, we will use measure theory in other areas of the book. For instance we will use $L^p$ spaces as examples of Banach spaces and their duals (cfr. Section 5.6 and its many mentions throughout the text). Another instance is the Spectral Theorem in its different incarnations (cfr. Section 12.4), which also requires measure and integration.

## 2.1. Motivation

Our first goal is to try and provide a convincing argument showing the need for a theory of measure and integration. We briefly discussed the Riemann integral on Section 1.3. In particular we know from Exercise 1.3.8 that given a continuous function $f$ on an interval $[a, b]$,

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{k=1}^{n} f\left(a + \frac{k(b-a)}{n}\right) \Delta_k, \tag{2.1}$$

where $\Delta_k = \frac{1}{n}$ for all $k$. More generally one can define the Riemann Integral in terms of upper and lower integrals (as we did on Definition 1.3.2), or in terms of Riemann sums over all kinds of partitions that shrink with $n$. These latter definitions allow one to have Riemann integrals of certain non-continuous functions. Actually, if $f$ has finitely many jump discontinuities, then (2.1) works with no glitches to produce a number that we call $\int_a^b f(t)\,dt$ (Exercise 1.3.8).

The idea in (2.1) is fundamental in Calculus and it is a key reason for the incredible success of Calculus in all areas of science. One of the first—if not the very first—motivations for (2.1) is, back to Newton, to find the distance travelled if we know the velocity $f(t)$ at time $t$. In that case, what happens in (2.1) can be described as taking small time-intervals, multiplying the (approximate) velocity on the interval $\Delta_k$ times the time-length of $\Delta_k$ (distance is velocity times time, when the velocity is constant), and adding; the result will approximate the distance travelled between the times $t = a$ and $t = b$. On the limit, we expect to obtain the exact distance travelled. With the same argument as above, we could think instead that $f(t)$ describes the acceleration, and now $\int_a^b f(t)\,dt$ represents the increase in velocity from $t = a$ to $t = b$. If we think of $f$ as a non-negative function, then $\int_a^b f(x)\,dx$ represents the area between the graph of $f$ and the $x$-axis, between $x = a$ and $x = b$. If $\rho(x)$ represents the linear density of an object, then $M = \int_a^b \rho(x)\,dx$ gives the mass; and $\frac{1}{M}\int_a^b x\,\rho(x)\,dx$ gives the centre of mass. There are of course several-variable versions of this, that allow one to use these ideas on the plane, or in space; and there are lots and lots of applications to every bit of science that uses mathematics one way or another.

Let us focus more on what the integral is. What (2.1) gives us is a recipe to produce a number, that we denote by $\int_a^b f(t)\,dt$, provided a good enough function $f$ and an interval $[a, b]$. What the recipe makes us do is, to take small parts $E_k$ of $[a, b]$ (say, $E_k$ could be the interval $\big[a + \frac{k(b-a)}{n}, a + \frac{(k+1)(b-a)}{n}\big]$); then multiply "value of $f$ in $E_k$ times size of $E_k$", and add. That is, the integral of $f$ over $[a, b]$ is the limit of sums of the form "value of $f$ in a region times the size of the region" for regions that subdivide $[a, b]$, as the sizes of the regions approach zero. So, if $E \subset \mathbb{R}$ is an interval we can rewrite (2.1) as

$$\int_E f(t)\,dt = \lim_{n\to\infty} \sum_{k=1}^{n} f(t_k^{(n)})\, m(E_k^{(n)}), \tag{2.2}$$

where $E_1^{(n)}, \dots, E_n^{(n)}$ are disjoint intervals that cover $E$, $t_k^{(n)} \in E_k^{(n)}$ for all $k = 1, \dots, n$, and $m(E_k^{(n)})$ is the length of $E_k^{(n)}$. Here are a few observations about (2.2):

- By rephrasing the definition as in (2.2), it works with no change when the domain is in $\mathbb{R}^m$ instead of in $\mathbb{R}$. We just need to use a different "size-function" $m$.
- No limiting occurs in the domain of $f$. In fact, all we need in the domain of $f$ is to be able to partition it, and to have a notion of size for these parts that allows us to choose small sizes.

- The limiting process occurs in the codomain of $f$. In fact, the integral is a limit of linear combinations of certain values of $f$. So we definitely need the codomain of $f$ to be a vector space, and with a topology that allows us to take limits; *and nothing else.* We do not multiply elements of the codomain. This means that we could take a continuous function $f : E \to \mathcal{X}$, with $E \subset \mathbb{R}^m$ and $\mathcal{X}$ any normed space, and then definition (2.2) works without any change whatsoever. Even the proof that the limit exists works verbatim, using the norm of $\mathcal{X}$ instead of the absolute value. Those not yet familiar with these terms can think that $\mathcal{X} = \mathbb{R}^n$ and instead of the absolute value the Euclidean norm is used.

So, all is fine, and we can actually generalize the integral to a fairly more general setting. Time to close this chapter... Not so fast! There are clouds obscuring the Riemannian sky.

For a subset $F$ of a set $E$, the **characteristic function** of $E$ is the function

$$1_E(t) = \begin{cases} 1, & t \in E \\ 0, & t \notin E \end{cases}$$

Consider the interval $[0,3]$ and the function $f : [0,3] \to \mathbb{R}$ given by $f(t) = 1_{[0,2]}$. This function has a single jump discontinuity, so its Riemann integral poses no problem:

$$\int_{[0,3]} 1_{[0,2]}(t)\,dt = \int_{[0,2]} 1\,dt + \int_{[2,3]} 0\,dt = 2.$$

As expected, we get that the integral is "value of the function" (here, 1 on $[0,2]$) "times size of the region" (here, 2). Of course the same will happen for any other interval.

Let $E = \mathbb{Q} \cap [0,1] \subset [0,2]$. If we want to calculate $\int_{[0,2]} 1_E(t)\,dt$, we should be able to do it without appealing to (2.1) and again expect to get "value of the function times size of the region". Now, what would the size of $E$ be? Fix $\varepsilon > 0$. Since $E$ is countable, we may write it as a sequence $E = \{q_k\}_{k\in\mathbb{N}}$. Then

$$E \subset \bigcup_{k\in\mathbb{N}} \Big(q_k - \frac{\varepsilon}{2^{k+1}}, q_k + \frac{\varepsilon}{2^{k+1}}\Big).$$

For any meaningful notion of length, we should expect that the larger set has the bigger length. And that the length of a union should be less than the lengths of the components of the union (as there could be overlap). So we expect, using again $m$ to denote the length,

$$m(E) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

As we can do this for any $\varepsilon > 0$, it has to be that $m(E) = 0$. We should expect, via "value of the function times size of the region" that $\int_{[0,2]} 1_E(t)\,dt = 0$. But the Riemann integral of $1_E$ does not exist; depending on the choices of $t_k^{(n)}$ in

(2.2), we can make the limit take any value in $[0,1]$, or we can make it fail to exist. This is a problem: if we know the (constant) value of the function and the size of the region, why would we have to say that the integral does not exist? Definition (2.2) seems nice and natural, but it is not that good for non-continuous functions.

The function $1_E$ highlights another shortcoming of the Riemann integral. If again we enumerate $\mathbb{Q}\cap[0,1]=\{q_k\}$, we can consider the functions $f_n(t)=1_{\{q_1,\ldots,q_n\}}(t)$. These functions have finitely many discontinuities, so they are Riemann integrable, and $\int_0^1 f_n(t)\,dt=0$ for all $n$. The functions $f_n$ are uniformly bounded and they converge to $1_E$ pointwise; but the integrals cannot converge to $\int_0^1 1_E$ because $1_E$ is not integrable.

What could we do instead of the Riemann Integral? We cannot leave the basic premise than an integral is (limit of) "sums of value of the function times size of the region", as that is the integral's essence; it is what makes integrals appear naturally in so many applications. One idea we could try is to subdivide the codomain, instead of the domain. How would this work? Suppose for notational simplicity that we have $f:[0,1]\to[0,1]$. We could partition the codomain in intervals $[y_{k-1},y_k]$; the "size of the region" where $f$ takes those values would be the length of $f^{-1}([y_{k-1},y_k])$. So, still denoting by $m$ the length function, we would have

$$\int_{[0,1]} f\,dm=\lim_{n\to\infty}\sum_{k=1}^n y_k\,m(f^{-1}([y_{k-1},y_k])). \tag{2.3}$$

The idea in (2.3) will lead us to define the Lebesgue Integral. But it will require us to address a few problems:

- As soon as $f$ stops being continuous, the pre-images of intervals are no longer going to be (unions of) intervals. So it is not entirely clear what the "length" function $m$ should be. At this point we will stop calling it length and start using the word "measure". We are looking for a function $m:\mathcal{P}(\mathbb{R})\to[0,\infty]$ that defines the measure of any subset of $\mathbb{R}$. A few natural properties that $m$ should satisfy are
  (1) $m([a,b])=b-a$ for all $a<b$; $m(\varnothing)=0$.
  (2) $m(X\cup Y)=m(X)+m(Y)$ if $X,Y$ are disjoint. More generally, $m(\bigcup_{k=1}^\infty E_k)=\sum_{k=1}^\infty m(E_k)$ if $E_1,E_2,\ldots,$ are disjoint. Note that we cannot expect the equality to work for uncountable unions, as we expect points to have measure zero and sets with nonzero measure are unions of points.
  (3) $m(s+X)=m(X)$ for any $s\in\mathbb{R}$. That is, $m$ should be translation invariant.

  Thing is, we will eventually show that no such $m$ as above exists; the family of all possible subsets of $\mathbb{R}$ is too crazy—we will prove this soon for $\mathbb{R}$ (Remark 2.3.18); for the case of $\mathbb{R}^3$, interested readers can check the Banach–Tarski Paradox on Chapter D, though it might not be an easy read. Banach and Tarski proved that the unit ball in $\mathbb{R}^3$ can

be divided into a finite number of pieces (as few as five) such that by appropriate rotations and translations they can be arranged as two copies of the unit ball. For such pieces, no reasonable notion of volume can be expected. Time to drop the towel? No. Smart people noticed that we can define $m$ on *some* subsets of $\mathbb{R}$, so that the above $m$ can be defined, and the subsets are plenty enough to cover most cases of interest.
- Once we restrict the subsets where $m$ makes sense, we run into problems with (2.3), where we want to evaluate $m$ on preimages of intervals. So not every function will work, and we will have to restrict to functions $f$ where the preimages are nice enough. The set of such functions, though, will turn out to be strictly bigger than the set of Riemann-integrable functions.

Addressing these issues is what this chapter is about. Namely, our program is to construct the measure $m$, and then use it to define an integral in the spirit of (2.3). We will end up obtaining way more than we bargained for, as Measure Theory works well beyond the motivational example from above, and the Lebesgue integral is notably more general than the Riemann integral.

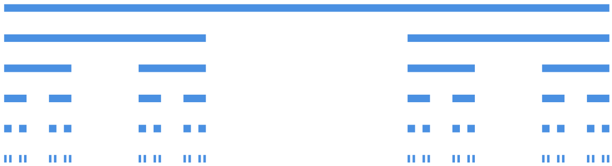
### 2.1.1. Exercises.

(2.1.1) Show that the integral in (2.1) exists if $f$ is continuous with the possible exception of finitely many jump discontinuities.

## 2.2. The Cantor Set

This section is not properly "measure theory". But, when someone is exposed to calculus and then basic analysis ("$\varepsilon$ and $\delta$"), the subsets of $\mathbb{R}$ that one considers are very mild; what we mean is that basically the only subsets of $\mathbb{R}$ that are considered are intervals. The Cantor set is not hard to construct, but it is far from trivial, and really really far from an interval.

The **Cantor set**, or more properly the **Ternary Cantor Set** is the set $\mathcal{C}$ obtained out of the closed interval $[0,1]$ by removing open "middle thirds". We start with $[0,1]$, and we remove $(\frac{1}{3},\frac{2}{3})$; then we remove $(\frac{1}{9},\frac{2}{9})$ and $(\frac{7}{9},\frac{8}{9})$, and so on. This would be a schematic of the first five steps:



The blue part is what remains after each step. As it is apparent from the picture, what remains after extracting the middle thirds seems rather thin. It is clearly nonempty, as for instance 0 and 1 will never be removed; actually, no endpoint of the middle thirds is ever removed, so $\mathcal{C}$ is infinite. In principle it is not clear how much of the original set remains after removing all middle thirds. The number of intervals removed duplicates on each step. If we denote the disjoint intervals removed in step $n$ by $C_{n,1},\ldots,C_{n,2^{n-1}}$, we can write

$$\mathcal{C}=[0,1]\setminus\bigcup_{n=1}^\infty\bigcup_{k=1}^{2^{n-1}}C_{n,k}.$$

Being the complement of an open set, $\mathcal{C}$ is closed. And, since $[0,1]$ is compact, $\mathcal{C}$ is compact. As all the intervals are disjoint, the length of $\mathcal{C}$ appears to be

$$m(\mathcal{C})=1-\sum_{n=1}^\infty\sum_{k=1}^{2^{n-1}}m(C_{n,k})=1-\sum_{n=1}^\infty\sum_{k=1}^{2^{n-1}}\frac{1}{3^n}=1-\sum_{n=1}^\infty\frac{2^{n-1}}{3^n}=1-\frac12\,\frac{2/3}{1-2/3}=0.$$

The above computation will be formally correct once we define Lebesgue measure; but it is not hard to be convinced that the intervals removed are pairwise disjoint, and that their total length adds to 1. So the Cantor set is what we call a nullset, it has length zero; in particular, it cannot contain any interval. We will see, though, that it is still a fairly significant set; the breakthrough is given by the next result. Recall that a number can be written in base 3; namely, any $t$ in $[0,1]$ can be written in a unique way as

$$t=\sum_{k=1}^\infty\frac{a_k}{3^k},\qquad a_k=\lfloor 3^k\,\big(t-\sum_{j=1}^{k-1}\frac{a_j}{3^j}\big)\rfloor \tag{2.4}$$

where $\{a_k\}\subset\{0,1,2\}$.

We will use the convention that if $s$ has a finite representation and its last digit is 1, we replace said 1 with $0222\cdots$; for instance, if $s=0.2011_3$, we write $s=0.2010222\cdots_3$ (this should not be scary if one looks at the analog choices in base 10; namely, one can choose to write $0.1999\cdots$ instead of $0.2$). With this convention, each $t\in[0,1]$ can be represented as in (2.4) in a unique way. For example,

$$0.2011_3=\frac23+\frac1{27}+\frac1{81}=\frac{54+3+1}{81}=\frac{58}{81},$$

but instead we use

$$0.2010222\cdots_3 = \frac{2}{3} + \frac{1}{27} + \sum_{k=5}^{\infty} \frac{2}{3^k} = \frac{19}{27} + \frac{2/3^5}{1-1/3} = \frac{19}{27} + \frac{1}{3^4} = \frac{58}{81}.$$

In base 2, the convention for uniqueness reduces to replacing any ending 1 with $0111\cdots$.

And this is why using base 3 is useful here:

**2.2.1. Proposition.**

*Let $\mathcal{C}$ be the Cantor ternary set. Then*

$$\mathcal{C} = \left\{ \sum_{k=1}^{\infty} \frac{a_k}{3^k} : \ \{a_k\} \subset \{0,2\}^{\mathbb{N}} \right\}.$$

*In other words, $\mathcal{C}$ consists of the numbers between 0 and 1 that can be expressed in base 3 without using the digit 1.*

*Proof.* One can show by induction that each $C_{n,k}$ is of the form $\left(\frac{3r+1}{3^n}, \frac{3r+2}{3^n}\right)$ for some integer $r$ (Exercise 2.2.1). Let $s = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \in C_{m+1,k}$. We have $C_{m+1,k} = (\ell/3^{m+1}, (\ell+1)/3^{m+1})$, with neither $\ell$ nor $\ell+1$ a multiple of 3, and

$$\ell < \sum_{j=1}^{m+1} \frac{3^{m+1}a_j}{3^j} + \sum_{j=m+2}^{\infty} \frac{3^{m+1}a_j}{3^j} < \ell + 1. \tag{2.5}$$

The sum in (2.5) is an integer, and

$$0 \le \sum_{j=m+2}^{\infty} \frac{3^{m+1}a_j}{3^j} \le 2 \sum_{j=m+2}^{\infty} \frac{1}{3^{j-m-1}} = 1. \tag{2.6}$$

If the series in (2.6) was equal to 1 we would have a contradiction in (2.5). So $\displaystyle\sum_{j=m+2}^{\infty} \frac{3^{m+1}a_j}{3^j} < 1$ and then

$$\ell = \sum_{j=1}^{m+1} \frac{3^{m+1}a_j}{3^j}.$$

If $a_{m+1} = 0$, then

$$\ell = 3^{m+1} \sum_{j=1}^{m} \frac{a_j}{3^j} \in 3\mathbb{N},$$

a contradiction. And if $a_{m+1} = 2$, then

$$\ell + 1 = 1 + \sum_{j=1}^{m+1} \frac{3^{m+1}a_j}{3^j} = 3 + \sum_{j=1}^{m} \frac{3^{m+1}a_j}{3^j} \in 3\mathbb{N},$$

again a contradiction. Thus $a_{m+1} = 1$. The converse works with the same steps from above. Namely, if $a_{m+1} \neq 1$ and $\ell$ is as in (2.5) (with possibly a $\leq$ on the left inequality), then the computations above show that $\ell \in 3\mathbb{N}$ or $\ell + 1 \in 3\mathbb{N}$, and then $s \notin C_{m+1,k}$ for any $k$. □

An informal proof of Proposition 2.2.1 amounts to notice that when we are removing the first middle third, we are removing all numbers of the form $0.1\cdots_3$. When we remove the next two middle thirds, we are removing the numbers of the form $0.01\cdots_3$ and those of the form $0.21\cdots_3$. Etc.

The main application of Proposition 2.2.1 is

***2.2.2. Proposition.***

*The Cantor set $\mathcal{C}$ is uncountable. More concretely, there exists a non-decreasing continuous surjection $\beta : \mathcal{C} \to [0,1]$, and $|\mathcal{C}| = |\mathbb{R}|$.*

*Proof.* Using the ternary representation from Proposition 2.2.1, define $\beta : \mathcal{C} \to [0,1]$ by

$$\beta\Big(\sum_{n=1}^{\infty} \frac{a_n}{3^n}\Big) = \sum_{n=1}^{\infty} \frac{a_n}{2^{n+1}}$$

Looking at the value of $\beta$ as $\sum_{n=1}^{\infty} \frac{a_n/2}{2^n}$, we can see it as the binary expansion of a number in $[0,1]$. We will show that $\beta$ is a continuous surjection.

The map $\beta$ is well-defined, since the ternary digits are unique (under the aforementioned assumption of replacing an ending 1 with $0222\cdots$). It is surjective: if $c \in [0,1]$, then we can write $c$ in base 2 by $c = \sum_{n=1}^{\infty} \frac{b_n}{2^n}$, with $b_n \in \{0,1\}$ for all $n$. Then

$$t = \sum_{n=1}^{\infty} \frac{2b_n}{3^n} \in \mathcal{C}$$

satisfies $\beta(t) = c$.

For continuity, if $|s - t| = \left|\sum_n \frac{a_n}{3^n} - \sum_n \frac{b_n}{3^n}\right| < \frac{1}{3^m}$ with $a_n, b_n \in \{0,2\}$, then $a_j = b_j$ for $j = 1, \ldots, m-1$ and so

$$|\beta(s) - \beta(t)| = \left|\sum_{n=m}^{\infty} \frac{a_n}{2^{n+1}} - \sum_{n=m}^{\infty} \frac{b_n}{2^{n+1}}\right| < \frac{1}{2^{m-1}}.$$

To see that $\beta$ is non-decreasing, we note that the usual order in $\mathbb{R}$ amounts in $[0,1]$ to the lexicographical order of the words $0.a_1a_2a_3\cdots$. This works both in base 2 and base 3, so $\beta$ preserves such order.

Since $[0,1]$ is the image of $\mathcal{C}$ under a surjection, we have $|\mathcal{C}| \geq |[0,1]| = |\mathbb{R}| \geq |\mathcal{C}|$. So Schröder–Bernstein guarantees that there exists a bijection $\mathcal{C} \to \mathbb{R}$. □

The use of Schröder–Bernstein in the proof of Proposition 2.2.2 is not gratuitous, since the map $\beta$ above is not injective: for instance, $\beta(1/3) = \beta(2/3)$. More fundamentally, since $\mathcal{C}$ is compact any continuous bijection $\mathcal{C} \to [0,1]$ would be a homeomorphism (Exercise 1.8.38), contradicting the fact that $[0,1]$ is connected while $\mathcal{C}$ is not.

The two-element pre-image of $1/2$ under $\beta$ corresponds to the two representations of $1/2$ as $0.1_2$ and $0.01111\cdots_2$; that is, $0.2_3 = \frac{2}{3}$ and $0.022\cdots_3 = \frac{1}{3}$.

Another consequence of Proposition 2.2.1 is that $\mathcal{C}$ contains more than the endpoints of the removed middle thirds—as there are countably many of these. One explicit example of a number in $\mathcal{C}$ that is not an endpoint of a middle third is $t = \frac{1}{4}$. That it is not an endpoint is trivial, since all endpoints have denominators which are powers of three. That $\frac{1}{4} \in \mathcal{C}$ follows from the fact that

$$\frac{1}{4} = 0.020202\cdots_3\,.$$

*2.2.3. Remark*. Because the endpoints of the middle thirds are not removed when constructing the Cantor set, it follows that $\mathcal{C}$ has no isolated points (Exercise 2.2.6). Then Corollary 1.8.22 gives another proof that $\mathcal{C}$ is uncountable. The Cantor set $\mathcal{C}$ also fails to contain any intervals (Exercise 2.2.7).

We will use the function $\beta$ to define the **Cantor Function** $\alpha : [0,1] \to [0,1]$. Namely,

$$\alpha(t) = \begin{cases} \beta(t), & t \in \mathcal{C} \\ \sup\{\beta(s) :\ s \in \mathcal{C} \cap [0,t]\}, & t \notin \mathcal{C} \end{cases}$$

It is not hard to see what $\alpha$ is doing. When $t \notin \mathcal{C}$, it belongs to one of the middle thirds $C_{n,k} = (a_{n,k}, b_{n,k})$ that was removed to construct $\mathcal{C}$; then we take $\alpha(t) = \beta(a_{n,k})$. In particular, $\alpha$ is constant in each $C_{n,k}$. As the length of $\mathcal{C}$ is 0 this means that $\alpha$—in terms that will be defined soon—is constant almost everywhere.

As $\beta$ is surjective, we automatically get that $\alpha$ is surjective. The previous paragraph implies that $\alpha$ is continuous almost everywhere, but in fact it is continuous everywhere. To see the continuity of $\alpha$ we consider two cases:

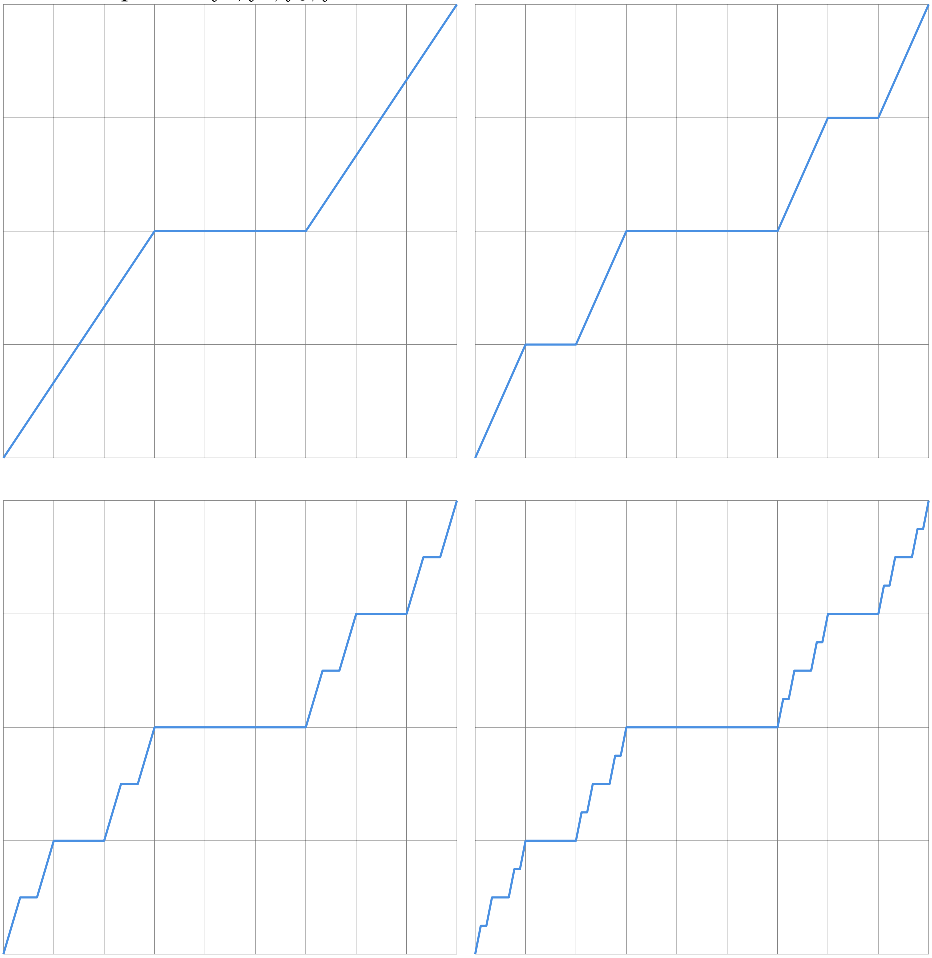
- $t \in [0,1] \setminus \mathcal{C}$. Say $t \in C_{n,k}$. As this is an open set, there exists $\delta > 0$ such that $|r - t| < \delta$ implies $r \in C_{n,k}$; then $|\alpha(t) - \alpha(r)| = 0$.
- $t \in \mathcal{C}$. Fix $\varepsilon > 0$. By the continuity of $\beta$, there exists $\delta > 0$ with $|\beta(t) - \beta(r)| < \varepsilon$ whenever $|t - r| < \delta$ and $r \in \mathcal{C}$. If $r \notin \mathcal{C}$ and $|t - r| < \delta$, we have that $r \in C_{n,k}$ for some $n, k$; let $r_1$ be the endpoint of $C_{n,k}$ closest to $t$. Then $|t - r_1| < |t - r| < \delta$ and $\alpha(r) = \beta(r_1)$. Thus
$$|\alpha(t) - \alpha(r)| = |\beta(t) - \beta(r_1)| < \varepsilon.$$

The function $\alpha$ maps $[0,1]$ onto $[0,1]$ in a continuous, non-decreasing way; in spite of the fact that its derivative is zero almost everywhere.

Alternatively, one can define $\alpha(x)$ as the (uniform) limit of the sequence of functions given by $f_0(x) = x$,

$$f_{n+1}(x) = \begin{cases} \frac{1}{2} f_n(3x), & x \in \left[0, \frac{1}{3}\right] \\ \frac{1}{2}, & x \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{1}{2} + \frac{1}{2} f_n(3x-2), & x \in \left[\frac{2}{3}, 1\right] \end{cases}$$

Below we picture $f_1, f_2, f_3, f_4$.



Yet another way to define $\alpha$ is the following: given $t = \sum_{k=1}^{\infty} \frac{a_k}{3^k}$, define

$$\alpha(t) = \begin{cases} \sum_{k=1}^{\infty} \frac{a_k}{2^{k+1}}, & \text{if } t \in \mathcal{C} \\ \sum_{k=1}^{\infty} \frac{b_k}{2^{k+1}}, & \text{if } t \notin \mathcal{C} \end{cases}$$

where $b_k$ is defined by letting $r = \min\{k : \ a_k = 1\}$, and

$$b_k = \begin{cases} a_k, & 1 \leq k \leq r-1 \\ 2, & k = r \\ 0, & k > r \end{cases}$$

The formulas above hide a bit the actual procedure, which is really simple. What $\alpha$ does is, for $t \in \mathcal{C}$, $\alpha(t) = \beta(t)$. And for $t \in [0,1] \setminus \mathcal{C}$, what $\alpha$ does is truncate at the first 1 in the ternary expansion, and then replace all the 2 with 1 and think of the result in base 2. For example if $t = 0.020212012_3$, then we truncate to $0.02021_3$ and thus $\alpha(t) = 0.01011_2$. In fractions, this says that

$$\alpha\Big(\frac{5000}{19683}\Big) = \frac{11}{32}.$$

We refer to Exercise 2.2.12 for a friendlier expression for $\alpha$ within this last approach.

For use in the coming exercises and elsewhere we define

**2.2.4. Definition.**

*The **dyadic numbers** are the numbers of the form $\dfrac{m}{2^k}$ with $m \in \mathbb{Z}$, $k \in \mathbb{N}$.*

### 2.2.1. Exercises.

(2.2.1) Show that for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^{n-1}\}$, the interval $C_{n,k}$ is of the form $\left(\frac{r}{3^n}, \frac{r+1}{3^n}\right)$, with neither $r$ nor $r+1$ multiples of 3.

(2.2.2) Let $t \in [0,2]$.

- (i) Show that there exist $a, b \in \mathcal{C}$ such that $t = a + b$.
- (ii) Find $a, b \in \mathcal{C}$, expressed as fractions, such that $a + b = 1$.
- (iii) Are such $a, b$ unique? If they are not, find another suitable pair $a, b$.

(2.2.3) Complete the details the proof of Proposition 2.2.2. That is, justify why if

$$|s - t| = \left|\sum_n \frac{a_n}{3^n} - \sum_n \frac{b_n}{3^n}\right| < \frac{1}{3^m}$$

with $a_n, b_n \in \{0, 2\}$, then $a_j = b_j$ for $j = 1, \ldots, m-1$.

(2.2.4) Show that the dyadic numbers in $[0,1]$ are dense.

(2.2.5) Consider the function $\beta$ from Proposition 2.2.2, and recall the notation $C_{n,k}$ for the removed intervals in the construction of $\mathcal{C}$.

(i) Show that the right endpoints of all the removed intervals $C_{n,k}$ are those numbers in $[0,1]$ such that their ternary expansion is finite, without any 1, and ends in 2.

(ii) Show that for any two endpoints of a removed interval $C_{n,k} = (a,b)$, we have $\beta(a) = \beta(b)$ and that this is a dyadic number.

(iii) Show that if $a, b \in \mathcal{C}$ are distinct and $\beta(a) = \beta(b)$, then there exist $n, k$ such that $C_{n,k} = (a,b)$.

(iv) Conclude that if $E$ is the set of endpoints of the removed intervals, then $\beta$ is injective on $\mathcal{C} \setminus E$.

(v) Conclude that $\beta^{-1}(\{t\})$ is a singleton if $t$ is not dyadic, and that it consists of two points when $t$ is dyadic.

(2.2.6) Show that $\mathcal{C}$ has no isolated points.

(2.2.7) Let $s, t \in \mathcal{C}$ with $s < t$. Show that there exists $b \in [0,1] \setminus \mathcal{C}$ with $s < b < t$. This shows that $\mathcal{C}$ is **totally disconnected**.

(2.2.8) Suppose that you create a set with a similar idea as the Cantor set, you start with the unit interval $[0,1]$ but instead of removing in each step middle intervals of measure $3^{-n}$, you remove middle intervals of measure $4^{-n}$. Discuss the set $\mathcal{D}$ you obtained. Does it contain any intervals? What properties are the same as in the Cantor set, and what properties are different?

(2.2.9) Define a function $f : [0,1] \to [0,1]$ in the following way. Write the complement of $\mathcal{D}$ (as in Exercise 2.2.8) as $\bigcup_n \bigcup_{k=1}^{2^n} D_{n,k}$ and the complement of $\mathcal{C}$ as $\bigcup_n \bigcup_{k=1}^{2^n} C_{n,k}$. For each $n, k$ let $f_{n,k}$ be the natural increasing bijection $D_{n,k} \to C_{n,k}$. That is, if $D_{n,k} = (a,b)$ and $C_{n,k} = (c,d)$, then $f_{n,k}(x) = c + \frac{(x-a)(d-c)}{b-a}$. Patch them together so that

$$f(x) = f_{n,k}(x), \qquad x \in D_{n,k}.$$

As the complement of the union of the $D_{n,k}$ is nowhere dense, we can extend $f$ by continuity to get $f : [0,1] \to [0,1]$. Show that $f$ is continuous, monotone non-decreasing, and fails the property that the preimage of a nullset is a nullset.

(2.2.10) Show that the sequence $\{f_n\}$ defined on page 88 converges uniformly to $\alpha$. This gives an alternative proof that $\alpha$ is continuous (and other properties, too).

(2.2.11) Consider the metric space

$$\mathcal{X} = \{f : [0,1] \to [0,1], \;\; \text{continuous}, \; f(0) = 0, \; f(1) = 1\},$$

with the metric

$$d(f,g)=\max\{|f(x)-g(x)|:\ x\in[0,1]\}.$$

Define $\Phi:\mathcal{X}\to\mathcal{X}$ by

$$(\Phi f)(x)=\begin{cases}\frac12\,f(3x), & x\in\left[0,\frac13\right]\\ \frac12, & x\in\left[\frac13,\frac23\right]\\ \frac12+\frac12\,f(3x-2), & x\in\left[\frac23,1\right]\end{cases}$$

We use $\Phi^n$ to denote composition of $\Phi$ with itself $n$ times.

(i) Show that $d(\Phi f,\Phi g)\leq\frac12\,d(f,g)$ for all $f,g\in\mathcal{X}$.

(ii) Show that, for any $f\in\mathcal{X}$, the sequence $\{\Phi^n f\}$ converges.

(iii) Show that, for any $f,g\in\mathcal{X}$, $\lim_n\Phi^n f=\lim_n\Phi^n g$.

(iv) Deduce that, for any $f\in\mathcal{X}$, $\lim_n\Phi^n f=\alpha$.

(2.2.12) Consider $\alpha$ as the fixed point in Exercise 2.2.11; that is do not use the other equivalent definitions.

(i) Let $x\in[0,1]$, written as $\sum_{k=1}^\infty\frac{a_k}{3^k}$. Put $m=\min\{k:\ a_k=1\}$ when $x\notin\mathcal{C}$, and $m=\infty$ when $x\in\mathcal{C}$. Show that

$$\alpha(x)=\sum_{k=1}^{m-1}\frac{a_k}{2^{k+1}}+\frac{1}{2^m}\qquad x\in[0,1]. \tag{2.7}$$

This works even when $m=\infty$, if we interpret $\frac{1}{2^\infty}=0$.

(ii) Show that $\alpha(1-x)=1-\alpha(x)$ for all $x\in[0,1]$.

(2.2.13) Let $\alpha:[0,1]\to[0,1]$ be the Cantor function. We saw that $\alpha$ is surjective and continuous. By Exercise 1.1.6 it admits a right-inverse. Show that such right-inverse cannot possibly be continuous.

## 2.3. Measures and Lebesgue Measure

Now we will get properly into measure theory. It is common to go either the "real route" (just do everything in $\mathbb{R}$) or the abstract route. We will go through a mixture. Lebesgue measure on the real line is needed. At the same time, for many results an abstract proof is no harder than a proof for the Lebesgue measure in the real line; hence the mixture. This will be clear as we go through.

If we want to measure lengths of sets in $\mathbb{R}$, some of those will be infinite. Because of that, and because some natural integrals are infinite as in the well-known improper Riemann integral

$$\int_0^1 \frac{1}{x}\,dx = \infty,$$

in measure theory it is common to work on the **extended real line**. What we do is include the two symbols $\infty$ and $-\infty$ to get a "closed interval" $[-\infty, \infty]$. The arithmetic of real numbers extends to these two symbols, with common sense rules like $a + \infty = \infty$, $a \times \infty = \infty$ if $a > 0$, $a \times \infty = -\infty$ if $a < 0$. We also take the slightly more controversial convention that $0 \times \infty = 0$. This makes sense in our context: a rectangle where one side is infinite and the other one is zero is a line, and we expect a line to have area zero.

**2.3.1. $\sigma$-Algebras.** As we mentioned when we discussed the shortcomings of the Riemann integral, we will eventually be able to define a measure as desired, but restricted to an appropriate family of sets. Towards that, we make the following definition.

**2.3.1. Definition.**

*Let $X$ be a set. A **$\sigma$-algebra** on $X$ is a family $\mathcal{A} \subset \mathcal{P}(X)$ such that*

*(i)* $\varnothing \in \mathcal{A}$;

*(ii)* $E \in \mathcal{A} \implies E^c \in \mathcal{A}$;

*(iii) if* $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, *then* $\bigcup_n E_n \in \mathcal{A}$.

*The pair $(X, \mathcal{A})$ is said to be a **measurable space**. The elements of $\mathcal{A}$ are the **measurable sets**.*

It is an essential feature of $\sigma$-algebras—and that is what the $\sigma$ stands for—that only countable unions have to stay in the algebra. Uncountable unions are not mentioned here; most of the time, we should expect them to possibly leave the algebra. Note that since the empty set is a valid choice, condition (iii) implies that $\mathcal{A}$ is closed under taking finite unions. Combined with (ii), both finite and countable intersections stay in $\mathcal{A}$.

As trivial examples of $\sigma$-algebras that work on any set $X$, consider $\mathcal{A} = \mathcal{P}(X)$ and $\mathcal{A} = \{\varnothing, X\}$. Most of the time, we will be looking at more interesting $\sigma$-algebras, although the $\sigma$-algebra $\mathcal{P}(X)$ will show up several times.

An artificial but slightly less obvious example of a $\sigma$-algebra is this: given a set $X$, let

$$\mathcal{A} = \{A \subset X : \ A \text{ is countable or } A^c \text{ is countable}\}.$$

It is not hard to show that $\mathcal{A}$ is a $\sigma$-algebra; in fact, it is the $\sigma$-algebra generated by the singletons $\big\{\{a\} : \ a \in X\big\}$ (Exercise 2.3.2), and it is strictly smaller than $\mathcal{P}(X)$ if and only if $X$ is uncountable.

***2.3.2. Proposition.***

*An arbitrary intersection of $\sigma$-algebras is a $\sigma$-algebra. In particular, given $S \subset \mathcal{P}(X)$, there exists a unique $\sigma$-algebra $\Sigma(S)$ such that $S \subset \Sigma(S)$ and if $S \subset \mathcal{A}$ for a $\sigma$-algebra $\mathcal{A}$, then $\Sigma(S) \subset \mathcal{A}$.*

*Proof.* Let $\{\mathcal{A}_j\}$ be a family of $\sigma$-algebras. Put $\mathcal{A} = \bigcap_j \mathcal{A}_j$. Since $\varnothing \in \mathcal{A}_j$ for all $j$, $\varnothing \in \mathcal{A}$. Similarly, $X \in \mathcal{A}$. If $A \in \mathcal{A}$ then $A \in \mathcal{A}_j$ for all $j$; so $A^c \in \mathcal{A}_j$ for all $j$, which means that $A^c \in \mathcal{A}$. Similarly, if $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, then $A_n \in \mathcal{A}_j$ for all $j$; this implies $\bigcup_n A_n \in \mathcal{A}_j$ for all $j$, and so $\bigcup_n A_n \in \mathcal{A}$.

Given $S \subset \mathcal{P}(X)$, let $\Sigma(S)$ be the intersection of all $\sigma$-algebras on $X$ that contain $S$ (such family is nonempty because $\mathcal{P}(X)$ is one such $\sigma$-algebra). By the above, $\Sigma(S)$ is a $\sigma$-algebra, and $S \subset \Sigma(S)$. And, by definition, if $\mathcal{A}$ is a $\sigma$-algebra with $S \subset \mathcal{A}$, then $\Sigma(S) \subset \mathcal{A}$. This last condition guarantees the uniqueness. $\square$

To talk about nontrivial concrete $\sigma$-algebras, let us go to the real line. Take $S \subset \mathcal{P}(\mathbb{R})$ to be

$$S = \{(a,\infty) : \ a \in \mathbb{R}\}.$$

By Proposition 2.3.2, there exists a smallest $\sigma$-algebra on $\mathbb{R}$ that contains $S$. We will call this $\sigma$-algebra the **Borel $\sigma$-algebra** and denote it by $\mathcal{B}(\mathbb{R})$. Our first observation will be that $\mathcal{B}(\mathbb{R})$ contains the open sets; and as $S$ is made of some open sets, this will show that $\mathcal{B}(\mathbb{R})$ is the $\sigma$-algebra generated by the open sets in $\mathbb{R}$.

Using that $\mathcal{B}(\mathbb{R})$ is a $\sigma$-algebra, we have that $(-\infty, a] = (a,\infty)^c \in \mathcal{B}(\mathbb{R})$ for all $a \in \mathbb{R}$. Then, for any $a \in \mathbb{R}$,

$$(-\infty, a) = \bigcup_{n=1}^\infty \Big( -\infty, a - \frac{1}{n}\Big] \in \mathcal{B}(\mathbb{R}).$$

Thus, for any $a, b \in \mathbb{R}$ with $a < b$,

$$(a,b) = (\infty, b) \cap (a,\infty) \in \mathcal{B}(\mathbb{R}).$$

Now Cantor's Lemma (Proposition 1.8.1) implies that all open sets are in $\mathcal{B}(\mathbb{R})$. And $\sigma$-algebras contain complements, so all closed sets are in $\mathcal{B}(\mathbb{R})$ too. And any countable intersection of open sets (that is, the $G_\delta$ sets ) is also there; as is any countable union of closed sets (the $F_\sigma$ sets). And countable unions and intersections of these, and complements, and so on. The singletons are in $\mathcal{B}(\mathbb{R})$, since they are closed; alternatively, we can write $\{a\} = \bigcap_n \Big(a - \dfrac{1}{n}, a + \dfrac{1}{n}\Big)$. It is hard to imagine that there is any subset of $\mathbb{R}$ that is not in $\mathcal{B}(\mathbb{R})$; but, as we will see later (Propositions 2.3.17 and 2.3.24), there are (many!) such subsets.

**2.3.3. Definition.**

*Let $X$ be a topological space with topology $\mathcal{T}$. The $\sigma$-algebra $\mathcal{B}(X)$ generated by $\mathcal{T}$ is the* ***Borel $\sigma$-algebra*** *for $X$ and is denoted by $\mathcal{B}(X)$.*

As in the case of $X = \mathbb{R}$, since $\mathcal{B}(X)$ is a $\sigma$-algebra and it contains all open sets, it also contains all closed sets, and all $G_\delta$ and $F_\sigma$ sets, etc.

**2.3.2. Measure and Outer Measure.** Again we will alternate between abstract and concrete.

**2.3.4. Definition.**

*Let $X$ be a set. If $\mathcal{A} \subset \mathcal{P}(X)$ is a $\sigma$-algebra, a* ***measure*** *on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that if $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ are pairwise disjoint, then*

$$\mu\Big(\bigcup_n A_n\Big) = \sum_n \mu(A_n) \qquad \textit{($\sigma$-additivity)} \tag{2.8}$$

*When $\mu$ is a measure on $\mathcal{A}$, we call the triple $(X, \mathcal{A}, \mu)$ a* ***measure space****.*

An immediate consequence of the definition is that $\mu(\varnothing) = 0$, by taking $A_n = \varnothing$ for all $n$, and then we get that (2.8) holds for finitely many disjoint $A_1, \dots, A_n$, since we can take $A_{n+1} = A_{n+2} = \cdots = \varnothing$.

Why only countably many pairwise disjoint sets in $\sigma$-additivity? Length on the real line provides an answer. If we define length from the starting point that the length of the interval $(a, b)$ is $b - a$, then points necessarily have length zero. But every set is a disjoint union of points, so if additivity held for arbitrary pairwise disjoint families of subsets, every set would have to have length zero.

*2.3.5. Example*. For an easy example of a measure space, consider any set $X$, $\mathcal{A} = \mathcal{P}(X)$, and $\mu(A) = |A|$ the cardinality of $A$ (without distinguishing infinities). This $\mu$ is known as the **counting measure** (Exercise 2.3.1).

*2.3.6. Example.* If $X$ is a set and $\mathcal{A}$ is the $\sigma$-algebra

$$\mathcal{A} = \{A \subset X : \quad A \text{ is countable or } A^c \text{ is countable}\},$$

define

$$\mu(A) = \begin{cases} 0, & A \text{ countable} \\ 1, & \text{otherwise} \end{cases}$$

Then $(X, \mathcal{A}, \mu)$ is a measure space (Exercise 2.3.3).

*2.3.7. Example.* Let $X$ be any set, and fix $x_0 \in X$. On $\mathcal{P}(X)$, define

$$\delta(S) = \begin{cases} 1, & x_0 \in S \\ 0, & x_0 \notin S \end{cases}$$

Then $(X, \mathcal{P}(X), \delta)$ is a measure space. The measure $\delta$ is an example of a **Dirac measure**. The set $\{x_0\}$ is an **atom** for $\delta$, as it is a singleton set with positive measure. A measure is **diffuse** if whenever $A$ is measurable and $\mu(A) > 0$, there exists $B \in \mathcal{A}$ with $B \subset A$ and $0 < \mu(B) < \mu(A)$.

At first sight it looks like the properties of a measure are rather scarce. At the very least, any notion of size should give a bigger number if we measure a bigger set. It turns out the definition does imply this, and several other properties too.

***2.3.8. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space. Then*

*(i) If $A, B \in \mathcal{A}$ and $A \subset B$, then*
$$\mu(A) \leq \mu(B).$$
*(ii) If $A, B \in \mathcal{A}$ and $A \subset B$, then*
$$\mu(B \setminus A) = \mu(B) - \mu(A).$$
*(iii) If $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$, then*
$$\mu\Big(\bigcup_n A_n\Big) \leq \sum_n \mu(A_n).$$
*(iv) if $\{E_k\} \subset \mathcal{A}$ is non-decreasing (that is, $E_k \subset E_{k+1}$ for all $k$) then*
$$\mu\Big(\bigcup_k E_k\Big) = \lim_{k \to \infty} \mu(E_k).$$
*(v) if $\{E_k\} \subset \mathcal{A}$ is non-increasing (that is, $E_k \supset E_{k+1}$ for all $k$) and if $\mu(E_k) < \infty$ for some $k$, then*
$$\mu\Big(\bigcap_k E_k\Big) = \lim_{k \to \infty} \mu(E_k).$$

*Proof.*

(i) We have
$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(ii) Looking at the equalities above,

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

(iii) Let $E_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k$. The sets $\{E_n\}$ are pairwise disjoint, they are in $\mathcal{A}$, and $\bigcup_n E_n = \bigcup_m A_n$. Then

$$\mu\Big(\bigcup_n A_n\Big) = \mu\Big(\bigcup_n E_n\Big) = \sum_n \mu(E_n) \leq \sum_n \mu(A_n).$$

(iv) If $\mu(E_n) = \infty$ for some $n$, then

$$\mu\Big(\bigcup_k E_k\Big) \geq \mu(E_n) = \infty.$$

And, as the sequence of numbers $\{\mu(E_k)\}$ is increasing, its limit is also $\infty$. So now we may suppose that $\mu(E_k) < \infty$ for all $k$. In this case,

$$\begin{aligned}\mu\Big(\bigcup_k E_k\Big) &= \mu\Big(E_1 \cup \bigcup_k E_{k+1} \setminus E_k\Big) = \mu(E_1) + \mu\Big(\bigcup_k E_{k+1} \setminus E_k\Big)\\ &= \mu(E_1) + \sum_k \mu(E_{k+1} \setminus E_k) = \mu(E_1) + \sum_k \mu(E_{k+1}) - \mu(E_k)\\ &= \lim_{k\to\infty} \mu(E_k).\end{aligned}$$

(v) Since $\mu(E_k) < \infty$ we have $\mu(E_n) < \infty$ for all $n \geq k$. So we may assume without loss of generality that $\mu(E_k) < \infty$ for all $k$. In that case, using what we have already proven in the non-decreasing case,

$$\begin{aligned}\mu\Big(\bigcap_k E_k\Big) &= \mu\Big(E_1 \setminus \bigcup_k (E_1 \setminus E_k)\Big) = \mu(E_1) - \mu\Big(\bigcup_k (E_1 \setminus E_k)\Big)\\ &= \mu(E_1) - \lim_{k\to\infty} \mu(E_1 \setminus E_k) = \lim_{k\to\infty} \mu(E_k).\end{aligned}$$

□

The last two properties in Proposition 2.3.8 are usually known as **continuity of the measure**. The name comes from the fact that the increasing union and the decreasing intersection can be thought of as $\lim_k E_k$, and then the aforementioned properties say that

$$\mu(\lim_k E_k) = \lim_k \mu(E_k).$$

In the continuity of the measure, it is possible to have a decreasing sequence $\{E_n\}$ of measurable sets in a measure space such that $\mu(E_n) = \infty$ for all $n$ but $\mu\big(\bigcap_n E_n\big) = 0$ (Exercise 2.3.11).

A slightly less obvious result is the following. The fact that a diffuse measure should "have every size in between two sizes" feels obvious if you think about it, but a formal proof is not that easy; hence the result below. It is here for reference, though, and it can be safely skipped .

***2.3.9. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a diffuse measure space where at least one nonempty set has finite measure. Then for any $A \in \mathcal{A}$ there exists a function $g : [0, \mu(A)] \to \mathcal{A}$ such that $\mu(g(t)) = t$, and $g(s) \subset g(t)$ whenever $s \leq t$.*

*Proof.* Let

$$\Gamma = \{(E, g) : \ E \subset [0, \mu(A)], \ g : E \to \mathcal{A}, \ g \text{ monotone}, \ \mu(g(t)) = t\}.$$

This set is trivially nonempty, as we can consider $E = \{\mu(A)\}$, $g(\mu(A)) = A$. We order $\Gamma$ by inclusion of graphs, i.e. $(E, g) \leq (F, h)$ if $E \subset F$ and $h|_E = g$. Let $\{(E_j, g_j)\}_{j \in J}$ be a chain in $\Gamma$. Define $E = \bigcup_j E_j \in \mathcal{A}$. For $t \in E$, there exists $j \in J$ with $t \in E_j$; define $g(t) = g_j(t)$. This is well-defined: if $E_j \subset E_k$, then $g_k(t) = g_j(t)$ from the monotony of the chain. So $(E_j, g_j) \leq (E, g)$ for all $j$, and the chain has an upper bound in $\Gamma$. By Zorn's Lemma, there exists a maximal $(E, h)$ in $\Gamma$.

If $E \subsetneq [0, \mu(A)]$, let $t_0 \in \overline{E}$. If $t_0 \notin E$, there exists a monotone sequence $\{t_n\} \subset E$ with $t_n \to t_0$. Suppose that $\{t_n\}$ is increasing. Let $E_0 = \bigcup_n h(t_n)$. By Proposition 2.3.8 we have that $\mu(E_0) = \lim_n h(t_n)$. As $h$ is monotone, we can define $h(t_0) = \mu(E_0)$; this would allow us to extend $h$ to the closure of $E$, contradicting the maximality of $(E, h)$. Similarly, when $t_n$ decreases to $t_0$, let $E_0 = \bigcap_n h(t_n)$. If $h_n(t_n) = \infty$ for all $n$ we put $h(t_0) = \infty$; otherwise, continuity of the measure applies and $\mu(E_0) = \lim_n h(t_n)$ and we can define $h(t_0) = E_0$ while keeping $h$ monotone. Again this contradicts the maximality of $(E, h)$, and thus $E$ is closed.

Knowing that $E$ is closed, its complement is open and is thus a countable union of disjoint intervals (Proposition 1.8.1). Suppose that $e_0 < e_1$ are in $E$, and $(e_0, e_1) \cap E = \varnothing$. As

$$\mu(h(e_1) \setminus h(e_0)) = \mu(h(e_1)) - \mu(h(e_0)) = e_1 - e_0 > 0,$$

since $\mu$ is diffuse there exists $A_0 \in \mathcal{A}$ with $A_0 \subsetneq h(e_1) \setminus h(e_0)$. Let $A_1 = h(e_0) \cup A_0$. Then $\mu(A_1) = e_0 + \mu(A_0) < e_1$. This allows us to extend $h$ to the point $t_1 = e_0 + \mu(A_0)$ by $\tilde{h}(t_1) = A_1$, again contradicting the maximality of $h$; so $E = [0, \mu(A)]$. □

In order to construct measures, the following notion will prove very useful.

**2.3.10. Definition.**

*An **outer measure** on $X$ is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that*

(i) $\mu^*(\varnothing) = 0$;

(ii) *if* $A \subset B$, *then* $\mu^*(A) \leq \mu^*(B)$;

(iii) *if* $\{S_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(X)$, *then*
$$\mu^*\Big(\bigcup_n S_n\Big) \leq \sum_n \mu^*(S_n)$$
*(countably subadditive/$\sigma$-subadditive).*

We emphasize that no $\sigma$-algebra is mentioned in Definition 2.3.10. The examples of measures mentioned above, where the $\sigma$-algebra is $\mathcal{P}(X)$, give basic examples of outer measures.

A more interesting example of an outer measure than those above occurs if we try to implement the program described in Section 2.1. The idea is to assign a value to the "length" of any subset of $\mathbb{R}$. Our starting point is that the length of an interval $(a, b)$ is $b - a$.

**2.3.11. Definition.**

*A **countable cover by intervals** for $S$ is a countable family $\mathcal{O} = \{I_k\}_{k\in\mathbb{N}}$ where each $I_k$ is an open interval and $S \subset \bigcup_k I_k$. Given a countable cover by intervals $\mathcal{O}$, its length is*
$$\ell(\mathcal{O}) = \sum_k \ell(I_k), \qquad \text{where } \ell((a, b)) = b - a.$$
*For $S \in \mathcal{P}(\mathbb{R})$ we define*
$$m^*(S) = \inf\Big\{\ell(\mathcal{O}) : \ \mathcal{O} \text{ is a countable cover by intervals for } S\Big\}. \tag{2.9}$$
*The function $m^*$ is called the **Lebesgue outer measure**.*

The idea behind $m^*$ follows fairly closely our intuition of how we would measure a length in real life. Namely, we would go through the continuous chunks of whatever we are measuring, and we would use a measuring tape between the end-points of each small chunk, getting a better approximation the smaller the chunks. It is tempting to think that the definition would work the same if we were to use finite instead countable covers; but that is not the case (Exercise 2.3.14).

To justify the name,

**2.3.12. Proposition.**

*The function $m^*$ is an outer measure on $\mathbb{R}$, and it satisfies $m^*(I) = \ell(I)$ for every interval $I$ (open or not). It is also **translation invariant**: $m^*(r+S) = m^*(S)$ for any $S \subset \mathbb{R}$ and $r \in \mathbb{R}$.*

*Proof.* We have $m^*(\varnothing) = 0$, since for every $\varepsilon > 0$ the cover $\mathcal{O}_\varepsilon = \{(0,\varepsilon)\}$ has $\ell(\mathcal{O}_\varepsilon) = \varepsilon$.

If $S \subset T$, then any countable open cover for $T$ is a countable open cover for $S$. Thus the infimum for $S$ is taken over possibly more sets, showing that $m^*(S) \le m^*(T)$.

Now consider a countable family $\{S_k\}$. If $m^*(S_k) = \infty$ for any $k$, then $m^*(\bigcup_k S_k) \le \sum_k m^*(S_k)$ holds trivially. So assume that $m^*(S_k) < \infty$ for all $k$. Fix $\varepsilon > 0$ and for each $k$ choose a countable cover by intervals $\{I_j^k\}_j$ such that

$$S_k \subset \bigcup_j I_j^k, \qquad \sum_j \ell(I_j^k) \le m^*(S_k) + \frac{\varepsilon}{2^k}.$$

Then $\{I_j^k\}_{j,k}$ is a countable cover by intervals for $\bigcup_k S_k$ and

$$m^*\Big(\bigcup_k S_k\Big) \le \sum_{k,j} \ell(I_j^k) \le \sum_k m^*(S_k) + \frac{\varepsilon}{2^k} = \varepsilon + \sum_k m^*(S_k).$$

As $\varepsilon$ was arbitrary, we obtain the $\sigma$-subadditivity.

For any interval $I = (a,b)$, we get by using the interval itself as its cover that $m^*(I) \le b-a$. On the other hand, for any countable cover by intervals $\mathcal{O} = \{I_k\}$ of $(a,b)$, and for any $n \in \mathbb{N}$ sufficiently large, the collection $\{I_k\}$ is an open cover of the compact set $\left[a+\frac{1}{n}, b-\frac{1}{n}\right]$. So there is a finite subcover with

$$\left[a+\frac{1}{n}, b-\frac{1}{n}\right] \subset I_{k_1} \cup \cdots \cup I_{k_m}.$$

If we write $I_{k_j} = (a_j, b_j)$ we may assume, by reordering so that $a_j \le a_{j+1}$ for all $j$ and removing any interval entirely contained in another one, that $b_j > a_{j+1}$ (if $b_j \le a_{j+1}$ the point $b_j \in (a,b)$ would not be covered). Then

$$\begin{aligned}\sum_{j=1}^m \ell(I_{k_j}) &= \sum_{j=1}^m b_j - a_j = b_m - a_1 + \sum_{j=1}^{m-1} b_j - a_{j+1}\\ &\ge b_m - a_1 \ge b - \frac{1}{n} - \Big(a + \frac{1}{n}\Big) = b - a - \frac{2}{n}.\end{aligned}$$

Doing this for every $n$ we conclude that $\ell(\mathcal{O}) \ge b-a$ for every countable open cover of $(a,b)$. Thus $m^*((a,b)) = b-a$. Now, for any $\varepsilon > 0$,

$$m^*([a,b]) \le m^*((a-\varepsilon, b+\varepsilon)) = b - a + 2\varepsilon.$$

Then $b-a = m^*((a,b)) \le m^*([a,b]) \le b-a$. Similarly we get that $m^*([a,b)) = b-a = m^*((a,b])$.

Finally, if $S \subset \mathbb{R}$ and $r \in \mathbb{R}$, for any countable open cover $\mathcal{O} = \{I_k\}$ of $S$, the family $r + \mathcal{O} = \{r + I_k\}$ is a countable open cover for $r + S$, and $\ell(\mathcal{O}) = \ell(r + \mathcal{O})$. This shows that $m^*(r + S) = m^*(S)$. □

*2.3.13. Remark*. An unintended consequence of Proposition 2.3.12 is that we get, after Theorem 1.6.22 and Corollary 1.8.22, a third proof of the uncountability of $\mathbb{R}$. Indeed, if $[0, 1]$ were countable then we would have $m^*([0, 1]) = 0$ with the same argument used for the rationals on page 81. This also shows that we cannot expect the proof of $m^*(I) \geq b - a$ above to be too elementary, for it has to fail if we use $\mathbb{Q}$ instead of $\mathbb{R}$. And it exhibits in a nice way how the order-completeness of $\mathbb{R}$ implies, via Heine–Borel, its uncountability.

**2.3.3. The Lebesgue $\sigma$-Algebra and the Lebesgue Measure.** As natural as the construction of $m^*$ feels, it is not a measure on $\mathcal{P}(\mathbb{R})$. Indeed, we will see that it is possible to construct a countable family $\{E_n\} \subset [0, 1]$, pairwise disjoint, with $m^*(E_n) = 1$ for all $n$; these have to be weird sets! We will provide details on this a little later (Remark 2.3.18), as we will go for the positive results first. What we will do is follow an amazing idea attributed to Carathéodory: it turns out that any outer measure becomes a measure on an appropriate $\sigma$-algebra.

***2.3.14. Theorem. (Carathéodory)***

*Let $X$ be a set and $\mu^*$ an outer measure on $X$. Then*

$$\mathcal{M}(X) = \big\{E \subset X : \text{ for all } S \subset X,\ \mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c)\big\}$$

*is a $\sigma$-algebra, and the restriction of $\mu^*$ to $\mathcal{M}(X)$ is a measure.*

*Proof.* The property

$$\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c)$$

is symmetric on $E$ and $E^c$; this shows that $\mathcal{M}(X)$ contains complements. We also have

$$\mu^*(S) = \mu^*(\varnothing) + \mu^*(S \cap X) = \mu^*(S \cap \varnothing) + \mu^*(S \cap \varnothing^c),$$

so $\varnothing, X \in \mathcal{M}(X)$. It only remains to deal with countable unions. Let us start humble, with unions of two sets. Given $E_1, E_2 \in \mathcal{M}(X)$ and $S \subset X$, we have

$$S \cap (E_1 \cup E_2) = (S \cap E_1) \cup (S \cap E_2) = (S \cap E_1) \cup (S \cap E_2 \cap E_1^c).$$

Using that $\mu^*$ is subadditive, we always have the inequality

$$\mu^*(S) \leq \mu^*(S \cap E) + \mu^*(S \cap E^c).$$

Then

$$\begin{aligned}\mu^*(S) &\le \mu^*(S\cap(E_1\cup E_2)) + \mu^*(S\cap(E_1\cup E_2)^c)\\ &= \mu^*((S\cap E_1)\cup(S\cap E_2\cap E_1^c)) + \mu^*(S\cap E_1^c\cap E_2^c)\\ &\le \mu^*(S\cap E_1) + \mu^*((S\cap E_1^c)\cap E_2) + \mu^*((S\cap E_1^c)\cap E_2^c)\\ &= \mu^*(S\cap E_1) + \mu^*(S\cap E_1^c) = \mu^*(S),\end{aligned}$$

where in the last two equalities we used first that $E_2 \in \mathcal{M}(X)$ and then that $E_1 \in \mathcal{M}(X)$. This shows that $E_1 \cup E_2 \in \mathcal{M}(X)$. Inductively we obtain that $E_1 \cup \cdots \cup E_n \in \mathcal{M}(X)$ whenever $E_1, \ldots, E_n \in \mathcal{M}(X)$.

As $\mathcal{M}(X)$ is closed under taking complements, we get that $\mathcal{M}(X)$ is closed under taking intersections of finitely many sets. It remains to show that we can go from finite unions to countable ones.

If $E_1, E_2 \in \mathcal{M}(X)$ with $E_1 \cap E_2 = \varnothing$, then

$$\begin{aligned}\mu^*(S\cap(E_1\cup E_2)) &= \mu^*(S\cap(E_1\cup E_2)\cap E_2)\\ &\qquad + \mu^*(S\cap(E_1\cup E_2)\cap E_2^c)\\ &= \mu^*(S\cap E_2) + \mu^*(S\cap E_1).\end{aligned}\tag{2.10}$$

Using the particular case where $S = X$ and induction, we have shown that $\mu^*$ is additive in $\mathcal{M}(X)$.

Let $\{A_k\}_{k\in\mathbb{N}} \subset \mathcal{M}(X)$. Put $E_0 = \varnothing$, and $E_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j$. Then the sets $\{E_n\}$ are pairwise disjoint and, since $\mathcal{M}(X)$ is closed under finite unions, finite intersections, and complements, $E_n \in \mathcal{M}(X)$ for all $n$. We also have $\bigcup_n E_n = \bigcup_k A_k$; call this set $E$. Our goal is to show that $E \in \mathcal{M}(X)$.

Let $F_n = \bigcup_{j=1}^{n} E_j \in \mathcal{M}(X)$. As $F_n \subset E$, we get that $E^c \subset F_n^c$. Then, for any $S \subset X$, using (2.10) inductively for the last equality,

$$\begin{aligned}\mu^*(S) &= \mu^*(S\cap F_n^c) + \mu^*(S\cap F_n) \ge \mu^*(S\cap E^c) + \mu^*(S\cap F_n)\\ &= \mu^*(S\cap E^c) + \sum_{k=1}^{n}\mu^*(S\cap E_k).\end{aligned}$$

As this can be done for any $n$, using that $\mu^*(S\cap E) = \mu^*\Big(\bigcup_k (S\cap E_k)\Big) \le \sum_k \mu^*(S\cap E_k)$,

$$\mu^*(S) \ge \mu^*(S\cap E^c) + \sum_{k=1}^{\infty}\mu^*(S\cap E_k) \ge \mu^*(S\cap E^c) + \mu^*(S\cap E) \ge \mu^*(S).$$

So $\mu^*(S) = \mu^*(S \cap E) + \mu^*(S \cap E^c)$. As we can do this for any $S$, we have shown that $E \in \mathcal{M}(X)$, and thus $\mathcal{M}(X)$ is a $\sigma$-algebra.

It remains to show that $\mu^*$ is a measure on $\mathcal{M}(X)$. And for that all we need to do is show that $\mu^*$ is $\sigma$-additive on $\mathcal{M}(X)$. So let $\{E_k\} \subset \mathcal{M}(X)$ be a countable family of pairwise disjoint sets. Then

$$\sum_{k=1}^{n} \mu^*(E_k) = \mu^*\Big(\bigcup_{k=1}^{n} E_k\Big) \leq \mu^*\Big(\bigcup_{k=1}^{\infty} E_k\Big) \leq \sum_{k=1}^{\infty} \mu^*(E_k).$$

As this works for all $n$, taking limit as $n \to \infty$ we get

$$\mu^*\Big(\bigcup_{k=1}^{\infty} E_k\Big) = \sum_{k=1}^{\infty} \mu^*(E_k).$$

□

When we start with an outer measure $\mu^*$, it is customary to denote the measure obtained via Theorem 2.3.14 by $\mu$.

The notation $\mathcal{M}(X)$ is not particularly good, as $\mathcal{M}(X)$ not only depends on the set $X$, but also on the outer measure. Still, in most cases this will be clear from context. In particular, by $\mathcal{M}(\mathbb{R})$ we mean that $\sigma$-algebra obtained via Theorem 2.3.14 from Lebesgue's outer measure $m^*$. The measure obtained as the restriction of $m^*$ to $\mathcal{M}(\mathbb{R})$ is called the **Lebesgue Measure** on $\mathbb{R}$ and it is denoted by $m$. It inherits from $m^*$ the property of being translation invariant:

***2.3.15. Proposition.***

*The Lebesgue $\sigma$-algebra and the Lebesgue measure are translation invariant.*

*Proof.* Since $m^*(r + E) = m^*(E)$, the equality will hold for $m$ as long as $r + E$ is measurable whenever $E$ is. So all we need to show is that $\mathcal{M}(\mathbb{R})$ is translation invariant. Let $S \subset \mathbb{R}$, $r \in \mathbb{R}$, and $E \in \mathcal{M}(\mathbb{R})$. Note that, for any $E \subset \mathbb{R}$, $(r + E)^c = r + E^c$. Then

$$\begin{aligned} m^*(S \cap (r + E)) &+ m^*(S \cap (r + E)^c) \\ &= m^*(r + [(S - r) \cap E]) + m^*(r + [(S - r) \cap E^c]) \\ &= m^*((S - r) \cap E) + m^*((S - r) \cap E^c) \\ &= m^*(S - r) = m^*(S). \end{aligned}$$

So $r + E \in \mathcal{M}(\mathbb{R})$. □

In principle, it could happen that $\mathcal{M}(X)$ is trivially small for certain outer measures. But that's not the case for our Lebesgue outer measure $m^*$:

***2.3.16. Theorem.***

*The $\sigma$-algebra $\mathcal{M}(\mathbb{R})$ contains the Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$.*

*Proof.* We want to show that $(a,\infty)\subset\mathcal{M}(\mathbb{R})$ for all $a\in\mathbb{R}$. This is enough, as then the $\sigma$-algebra $\mathcal{B}(\mathbb{R})$ generated by such sets will be contained in the $\sigma$-algebra $\mathcal{M}(\mathbb{R})$.

Let $a\in\mathbb{R}$, $S\subset\mathbb{R}$, and $\mathcal{O}=\{I_k\}$ a countable open cover of $S$. Fix $\varepsilon>0$ and define

$$\mathcal{O}_1=\{I_k\cap(a,\infty)\},\qquad \mathcal{O}_2=\Big\{I_k\cap\big(-\infty,a+\tfrac{\varepsilon}{2^k}\big)\Big\}.$$

These are, respectively, countable open covers of $S\cap(a,\infty)$ and $S\cap(-\infty,a]$. As $\ell((a,\infty)\cap\big(-\infty,a+\frac{\varepsilon}{2^k}\big))=\frac{\varepsilon}{2^k}$,

$$\ell(I_k\cap(a,\infty))+\ell\Big(I_k\cap\Big(-\infty,a+\frac{\varepsilon}{2^k}\Big)\Big)\le\ell(I_k)+\frac{\varepsilon}{2^k}$$

and then

$$\ell(\mathcal{O}_1)+\ell(\mathcal{O}_2)\le\ell(\mathcal{O})+\varepsilon.$$

We get

$$m^*(S\cap(a,\infty))+m^*(S\cap(-\infty,a])\le\ell(\mathcal{O}_1)+\ell(\mathcal{O}_2)\le\ell(\mathcal{O})+\varepsilon.$$

This can be done for any $\varepsilon>0$ and any countably open cover $\mathcal{O}$ for $S$, so

$$m^*(S\cap(a,\infty))+m^*(S\cap(-\infty,a])\le m^*(S).$$

The reverse inequality is free from $m^*$ being an outer measure, giving us equality. This says that $(a,\infty)\in\mathcal{M}(\mathbb{R})$. $\square$

As we have seen, when we have a measure space $(X,\mathcal{A},\mu)$, the sets in $\mathcal{A}$ are called **measurable**. In the particular case of the Lebesgue measure, the $\sigma$-algebra $\mathcal{M}(\mathbb{R})$ is the **Lebesgue $\sigma$-algebra** and its sets are **Lebesgue measurable**.

In the same way that in Calculus every function one naturally comes up with is at worst piecewise differentiable, it is far from trivial to construct non-measurable sets in the real line. But they do exist.

***2.3.17. Proposition. (Vitali)***

*The inclusion $\mathcal{M}(\mathbb{R})\subset\mathcal{P}(\mathbb{R})$ is strict.*

*Proof.* Define a relation, on $\mathbb{R}$, by $x\sim y$ if $x-y\in\mathbb{Q}$. It is routine to check that $\sim$ is an equivalence relation. Let $E$ be a set of representatives of $\sim$ in

$[0,1)$. Define an operation (you could call it *cyclic addition*)

$$a \dotplus b = \begin{cases} a+b, & \text{if } a+b<1 \\ a+b-1, & \text{if } a+b \geq 1 \end{cases}$$

For each $r \in \mathbb{Q} \cap [0,1)$, let

$$E_r = \{x \dotplus r : \ x \in E\} \subset [0,1].$$

Then

(1) if $r \neq s$, then $E_r \cap E_s = \varnothing$.

Indeed, if $z \in E_r \cap E_s$ there exist $x, y \in E$ such that $z = x \dotplus r = y \dotplus s$. Then, whether there is a "minus one" or not on each side of the equality, we get that $x - y \in \mathbb{Q}$, which means $y = x$ (as they are representatives). Then, cancelling $x$ from both sides on $x \dotplus r = x \dotplus s$, we get that either $r = s$ or $r = s \pm 1$, the latter being impossible. So $r = s$ and $E_r = E_s$.

(2) $\displaystyle\bigcup_{r \in \mathbb{Q}\cap[0,1)} E_r = [0,1)$.

Fix $x \in [0,1)$. By definition of $E$, there exists $y \in E$ with $x \sim y$. If $y \leq x$, let $r = x - y \in \mathbb{Q} \cap [0,1)$. Then $x = y + r \in E_r$. If, on the other hand, $x < y$, let $r = 1 - (y - x) \in \mathbb{Q} \cap [0,1)$; as $y + r = 1 + x > 1$, we get that

$$x = y + r - 1 = y \dotplus r \in E_r.$$

(3) If $E$ is Lebesgue measurable, then $E_r$ is Lebesgue measurable and, for any $r \in \mathbb{Q} \cap [0,1)$, $m(E_r) = m(E)$.

Suppose that $E \in \mathcal{M}(\mathbb{R})$. We have

$$\begin{aligned} E_r &= \{x + r : \ x \in E \cap [0, 1-r)\} \cup \{x + r - 1 : \ x \in E \cap [1-r, 1)\} \\ &= (r + E \cap [0, 1-r)) \cup (r - 1 + E \cap [1-r, 1)). \end{aligned} \tag{2.11}$$

Both sets in the union are translates of measurable sets, so measurable by Proposition 2.3.15. Thus $E_r \in \mathcal{M}(\mathbb{R})$. Then

$$\begin{aligned} m(E_r) &= m(r + E \cap [0,1-r)) + m(r - 1 + E \cap [1-r,1)) \\ &= m(E \cap [0,1-r)) + m(E \cap [1-r,1]) = m(E). \end{aligned}$$

So $m(E_r) = m(E)$ for all $r$.

The above properties imply that $E$ cannot be Lebesgue measurable. If it were, then by (2.11) each $E_r$ would be measurable, and

$$1 = m([0,1)) = m\Big(\bigcup_r E_r\Big) = \sum_r m(E_r) = \sum_r m(E).$$

The equality is contradictory both when $m(E) = 0$ and when $m(E) > 0$. Thus $E \notin \mathcal{M}(\mathbb{R})$. $\square$

The set $E$ in the proof of Proposition 2.3.17 is known as the **Vitali set**. It is not properly "the" Vitali set, though, as we made an implicit choice when we chose the set of representatives. What is always true, in light of Lemma 2.3.20 below, is that $m^*(E) > 0$ for any choice of the representatives. Now, because the classes are invariant by shifting by a rational, we could fix any interval $(a, b) \subset [0, 1]$ and choose the representatives inside $(a, b)$. This means in particular that we can construct a Vitali set $E$ with $m^*(E) < \varepsilon$ for any fixed $\varepsilon > 0$. On the other hand, there is no upper bound on the possible outer measure of the Vitali set; see Remark 2.3.18.

In any case, Proposition 2.3.17 constructs a countable, pairwise disjoint family $\{E_r\}_{r\in\mathbb{Q}}$ of subsets of $[0, 1]$, each with equal positive Lebesgue outer measure. But things can get weirder.

*2.3.18. Remark*. Appealing to a bit of deeper set theory, an argument we learned from Robert Israel shows that it is possible to construct Vitali sets of arbitrary outer measure. Let $\mathcal{F} = \{W \subset [0, 1],\ \text{open}, m(W) < 1\}$ (note that there are uncountably many open subsets of $[0, 1]$ with Lebesgue measure 1). We have $|\mathcal{F}| = |\mathbb{R}|$ (because each open set which is not an interval is at most a countable union of intervals, recall Proposition 1.8.1, and because we can index the endpoints of intervals by $\mathbb{R}$). Then we can well-order $\mathcal{F}$ in such a way that for all $W \in \mathcal{F}$ we have $|\{U \in \mathcal{F} : U < W\}| < |\mathbb{R}|$ (this is analog to the situation where in the usual order in $\mathbb{N}$, we have $|\{n \in \mathbb{N} :\ n < m\}| < |\mathbb{N}|$). Because $m(W) < 1$ for all $W \in \mathcal{F}$, we always have that $|[0, 1] \setminus W| = |\mathbb{R}|$. This allows us to choose, for each $W \in \mathcal{F}$ an $f(W) \in [0, 1]$ such that $f(W) \notin W$ and also $f(W) \notin \{f(U) + q :\ U \in \mathcal{F},\ U < W,\ q \in \mathbb{Q}\}$. The reason is cardinality, as this latter set has cardinality strictly less than $|\mathbb{R}|$ (as guaranteed by the well-order) and we have $|\mathbb{R}|$ points in $[0, 1] \setminus W$ to choose from. By construction the set $\{f(W) :\ W \in \mathcal{F}\}$ contains at most one representative from each Vitali class. Now form $E$ by taking $\{f(W) :\ W \in \mathcal{F}\}$ and put representatives of the classes that were not there yet. This set $E$ is not Lebesgue measurable, as we can repeat the argument from the proof of Proposition 2.3.17. As for $m^*(E)$, if $E \subset U$ for some open set $U$, by construction we have that if $m(U) < 1$ then $f(U) \notin U$; thus $m(U) = 1$. This shows that $m^*(E) = 1$.

Using the Vitali set from the previous paragraph we can fulfill what we mentioned on page 100. If we repeat the proof of Proposition 2.3.17 with this Vitali set $E$, we get a countable family $\{E_r\}_{r\in\mathbb{Q}\cap[0,1]}$ of pairwise disjoint subsets of $[0, 1]$ and with $m^*(E_r) = m^*(E) = 1$ for all $r$. In particular, if we take $r \neq s$ we get $E_r \cap E_s = \varnothing$, $m^*(E_r) = m^*(E_s) = 1$ and $m^*(E_r \cup E_s) = 1 < 2 = m^*(E_r) + m^*(E_s)$ and the additivity fails. This is in some sense worrying, as the "natural" idea we used to construct $m^*$ allows us to decompose $[0, 1]$ into countably many disjoint subsets, each of length one. But we should not see this as a failure of the idea in the definition of $m^*$, but rather as evidence of how pathological arbitrary subsets of the real line can be. When we stay with the "good" subsets of $\mathbb{R}$, namely those in $\mathcal{M}(X)$, Lebesgue measure works wonderfully. And $\mathcal{M}(X) \supset \mathcal{B}(\mathbb{R})$, so the Lebesgue $\sigma$-algebra is plenty big.

What about the inclusion $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$? This is also proper, as we will see. We just need a bit of preparation. In a measure space $(X, \mathcal{A}, \mu)$, measurable nullsets are in a sense irrelevant—after all, they are precisely the sets with "size zero". This leads to the following terminology:

**2.3.19. Definition.**

*Given an outer measure $\mu^*$ on $X$, we say that $E \subset X$ is a **nullset** if $\mu^*(E) = 0$.*

*If a property holds for every point in $X \setminus A$, with $A$ a nullset, we say that said property holds **almost everywhere**, often denoted **a.e.** For instance when we say that $f = g$ a.e., what we mean is that there exists $A \subset X$ with $\mu(A) = 0$ and such that $f = g$ on $X \setminus A$.*

*If a measure space $(X, \mathcal{A}, \mu)$ has the property that all subsets of nullsets are measurable, we say that the measure $\mu$ is **complete**.*

It is important to keep in mind that most of the time talking about "almost everywhere" only makes sense on complete spaces.

***2.3.20. Lemma.***

*Let $\mu^*$ be an outer measure on $X$ and $E \subset X$. If $\mu^*(E) = 0$, then $E \in \mathcal{M}(X)$. That is, nullsets are measurable. In particular, the restriction of $\mu^*$ to $\mathcal{M}(X)$ is a complete measure.*

*Proof.* For any $S \subset X$,

$$\begin{aligned} \mu^*(S) &\leq \mu^*(S \cap E) + \mu^*(S \cap E^c) \leq \mu^*(E) + \mu^*(S \cap E^c) \\ &= \mu^*(S \cap E^c) \leq \mu^*(S). \end{aligned}$$

Thus $E \in \mathcal{M}(X)$. □

The situation in Lemma 2.3.20 should not confuse us in the case where a measure space $(X, \mathcal{A}, \mu)$ is given. When there is no outer measure involved, we are not guaranteed that subsets of a nullset are measurable. As a rather obvious example, we can have the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ where, as we will show, there are nullsets which are not Borel measurable. We will see more examples of this situation in Section 2.7. This is not much of an issue because we can always complete a measure:

***2.3.21. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space. Then*

$$\overline{\mathcal{A}} = \{A \cup B : \ A \in \mathcal{A}, \ B \subset X, \ \exists B' \in \mathcal{A}, \ B \subset B', \ \mu(B') = 0\}$$

*is a $\sigma$-algebra that contains $\mathcal{A}$, and*

$$\bar{\mu}(A \cup B) = \mu(A), \qquad\qquad A \in \overline{\mathcal{A}}$$

*defines a measure on $\overline{\mathcal{A}}$ that agrees with $\mu$ on $\mathcal{A}$ and such that $(X, \overline{\mathcal{A}}, \bar{\mu})$ is a complete measure space.*

*Proof.* In all the arguments below we use $A$ to denote an arbitrary element of $\mathcal{A}$, and $B$ to denote a subset of $X$ such that there exists $B' \in \mathcal{A}$ with $B \subset B'$ and $\mu(B') = 0$.

We check first that $\overline{\mathcal{A}}$ is a $\sigma$-algebra. If $\{A_n\} \subset \mathcal{A}$ and $\{B_n\}$ are subsets of $X$ such that there exist $\{B_n'\} \subset \mathcal{A}$ with $\mu(B_n') = 0$ for all $n$, then

$$\bigcup_n (A_n \cup B_n) = \bigcup_n A_n \cup \bigcup_n B_n \in \overline{\mathcal{A}},$$

since $\bigcup_n B_n \subset \bigcup_n B_n'$, $\bigcup_n B_n' \in \mathcal{A}$ and $\mu\Big(\bigcup_n B_n'\Big) \leq \sum_n \mu(B_n') = 0$. So $\overline{\mathcal{A}}$ contains countable unions of its elements. As for complements, if $A \cup B \in \overline{\mathcal{A}}$, then there exists $B' \in \mathcal{A}$ with $B \subset B'$ and $\mu(B') = 0$. We have

$$\begin{aligned}(A \cup B)^c = A^c \cap B^c &= A^c \cap \big((B')^c \cup (B' \setminus B)\big) \\ &= \big(A^c \cap (B')^c\big) \cup \big(A^c \cap (B' \setminus B)\big) \in \overline{\mathcal{A}},\end{aligned}$$

since $A^c \cap (B')^c \in \mathcal{A}$ and $A^c \cap (B' \setminus B) \subset B'$.

Before checking that $\bar{\mu}$ is a measure, we need to verify that it is well-defined. Suppose that $A \cup B = C \cup D$, with $A, C \in \mathcal{A}$ and there exist $B', D' \in \mathcal{A}$ nullsets with $B \subset B'$ and $D \subset D'$. Using that $B' = B \cup B'$ and $D' = D \cup D'$,

$$\begin{aligned}\mu(A) = \mu(A \cup B' \cup D') &= \mu(A \cup B \cup B' \cup D') = \mu(C \cup D \cup B' \cup D') \\ &= \mu(C \cup B' \cup D') = \mu(C).\end{aligned}$$

So $\bar{\mu}$ is well-defined. To check that it is a measure, we already have that $\bar{\mu}(\varnothing) = \mu(\varnothing) = 0$ since $\varnothing \in \mathcal{A}$. And if $\{A_n \cup B_n\}$ is a pairwise disjoint sequence in $\overline{\mathcal{A}}$ with $A_n \in \mathcal{A}$ and $B_n \subset B_n'$, $B_n' \in \mathcal{A}$, $\mu(B_n') = 0$ for all $n$,

$$\begin{aligned}\bar{\mu}\Big(\bigcup_n (A_n \cup B_n)\Big) = \bar{\mu}\Big(\bigcup_n A_n \cup \bigcup_n B_n\Big) &= \mu\Big(\bigcup_n A_n\Big) \\ &= \sum_n \mu(A_n) = \sum_n \bar{\mu}(A_n \cup B_n),\end{aligned}$$

since $\bigcup_n B_n$ is a subset of the measurable nullset $\bigcup_n B'_n$. □

*2.3.22. Remark* . In any measure space, a countable union of nullsets is a nullset. In the case of Lebesgue measure, any countable set is a nullset (this was justified on page 81); the Cantor set is an uncountable nullset.

Since we are talking measurability and lack thereof, this is probably a good point to mention the Banach–Tarski Paradox. This states the fact that it is possible to decompose a ball in $\mathbb{R}^3$ into a finite number of disjoint pieces, rotate a few of them, and reassemble then as two balls of the same size. For such sets we cannot expect a coherent notion of volume; that is, they have to be non-measurable. Details about the construction are provided in Chapter D.

### 2.3.4. The Borel and Lebesgue $\sigma$-Algebras.

***2.3.23. Lemma.***

*Let $E \in \mathcal{M}(\mathbb{R})$ with $m(E) > 0$. Then there exists $V \subset E$ with $V \notin \mathcal{M}(\mathbb{R})$.*

*Proof.* For each $k \in \mathbb{N}$, let $E_k = E \cap [-k, k]$. As $E = \bigcup_k E_k$ is an increasing union, by continuity of the measure (Proposition 2.3.8) there exists $k$ with $0 < m(E_k) < \infty$. As in Proposition 2.3.17, consider on $E_k$ the equivalence relation $x \sim y$ if $x - y \in \mathbb{Q}$. Let $V$ be a set of representatives in $E_k$, and number the countable set $\mathbb{Q}\cap[-2k, 2k]$ as $\{q_n\}$. For each representative $x \in V$ and for any $y \in E_k$ with $y \sim x$ we have $y - x \in \mathbb{Q} \cap [-2k, 2k] = \{q_n\}$. So its class $[x]$ satisfies $[x] \subset x + \{q_n\}$. This gives us

$$E_k = \bigcup_{x\in V} [x] \subset \bigcup_{x\in V} x + \{q_n\} = V + \{q_n\} = \bigcup_n (V + q_n).$$

Note that the sets $\{q_n+V\}_n$ are pairwise disjoint: for if $z \in (q_n+V)\cap(q_m+V)$, then $z = q_n + x = q_m + y$ for some $x, y \in V$; then $x - y = q_m - q_n \in \mathbb{Q}$, so $x = y$ as they are representatives and then $q_n = q_m$.

If $V \in \mathcal{M}(\mathbb{R})$, then also $q_n + V \in \mathcal{M}(\mathbb{R})$ by Proposition 2.3.15. For any $j \in \mathbb{N}$, as $V + q_n \subset [-3k, 3k]$,

$$6k = m([-3k, 3k]) \geq m\Big(\bigcup_{n=1}^{j} (q_n + V)\Big) = \sum_{n=1}^{j} m(q_n + V) = j\, m(V).$$

Since we can do this for all $j$, this forces $m(V) = 0$. But then
$$0 < m(E_k) \leq m\Big(\bigcup_{n=1}^{\infty}(q_n + V)\Big) = \sum_{k=1}^{\infty} m(q_n + V) = 0$$
gives us a contradiction. This means that $V \notin \mathcal{M}(\mathbb{R})$, while $V \subset E$. $\square$

***2.3.24. Proposition.***

*The inclusion $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$ is strict.*

*Proof.* Consider the Cantor Ternary function $\alpha : [0,1] \to [0,1]$ as in Section 2.2. Extend it to all of $\mathbb{R}$ by
$$\tilde{\alpha}(t) = \begin{cases} 0, & \text{if } t < 0 \\ \alpha(t), & \text{if } t \in [0,1] \\ 1, & \text{if } t > 1 \end{cases}$$
and define $\gamma(t) = \tilde{\alpha}(t) + t$. This function is continuous and bijective. Define
$$\mathcal{S} = \{\gamma^{-1}(B) : \; B \in \mathcal{B}(\mathbb{R})\}.$$
This is a $\sigma$-algebra, as preimaging preserves all set operations. From $\gamma$ being continuous, surjective, increasing, and unbounded,
$$(a, \infty) = \gamma^{-1}\big((\gamma(a), \infty)\big) \in \mathcal{S}, \qquad\qquad a \in \mathbb{R}.$$
Thus $\mathcal{B}(\mathbb{R}) \subset \mathcal{S}$. As the Cantor set $\mathcal{C}$, being closed, is in $\mathcal{B}(\mathbb{R})$, we have $\mathcal{C} \in \mathcal{B}(\mathbb{R}) \subset \mathcal{S}$; So $\mathcal{C} = \gamma^{-1}(\mathcal{C}')$ for some $\mathcal{C}' \subset \mathcal{B}(\mathbb{R})$, and so $\gamma(\mathcal{C}) = \mathcal{C}' \in \mathcal{B}(\mathbb{R})$ (here we use that $\gamma$ is onto, see Exercise 1.1.5). Enumerating the removed middle thirds as $\{(a_k, b_k)\}_k$, we can write
$$\mathcal{C} = [0,1] \setminus \bigcup_k (a_k, b_k).$$
Now, using that the intervals $(a_k, b_k)$ are all disjoint and that $\tilde{\alpha}$ is constant on each of them so in particular $\tilde{\alpha}(a_k) = \tilde{\alpha}(b_k)$ for all $k$,
$$\begin{aligned} 2 &= m([0,2]) = m(\gamma([0,1])) = m(\gamma(\mathcal{C}) \cup \gamma([0,1] \setminus \mathcal{C})) \\ &= m(\gamma(\mathcal{C})) + \sum_k m(\gamma((a_k, b_k))) = m(\gamma(\mathcal{C})) + \sum_k m(\tilde{\alpha}(a_k) + a_k, \tilde{\alpha}(b_k) + b_k) \\ &= m(\gamma(\mathcal{C})) + \sum_k (b_k - a_k) = m(\gamma(\mathcal{C})) + m([0,1] \setminus \mathcal{C}) = m(\gamma(\mathcal{C})) + 1. \end{aligned}$$
Thus $m(\gamma(\mathcal{C})) = 1$. By Lemma 2.3.23 there exists $V \subset \gamma(\mathcal{C})$ non-measurable. From $\gamma^{-1}(V) \subset \mathcal{C}$ we have that $\gamma^{-1}(V)$ is a nullset, which is then measurable by Lemma 2.3.20. And $\gamma^{-1}(V) \notin \mathcal{B}(\mathbb{R})$. For if $\gamma^{-1}(V) \in \mathcal{B}(\mathbb{R}) \subset \mathcal{S}$ we would

have $V \in \mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, a contradiction since $V \notin \mathcal{M}(\mathbb{R})$. We have then established that $\gamma^{-1}(V) \in \mathcal{M}(\mathbb{R}) \setminus \mathcal{B}(\mathbb{R})$. ☐

The example of a Lebesgue-measurable set that is not Borel-measurable is a nullset. And that is basically it. As we will see, every Lebesgue-measurable set differs from a Borel set by a nullset. One can prove Proposition 2.3.24 in a more brutal way, in the sense that $|\mathcal{B}(\mathbb{R})| < |\mathcal{M}(\mathbb{R})|$ in cardinality terms. Using a bit of non-elementary set theory (ordinals) it is possible to show that $|\mathcal{B}(\mathbb{R})| = |\mathbb{R}|$. On the other hand, we know that the Cantor set $\mathcal{C}$ is closed, so $\mathcal{C} \in \mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$ (but we do not need the fact that it is in $\mathcal{B}(\mathbb{R})$ here). As shown in Lemma 2.3.20, every subset of $\mathcal{C}$, being a nullset, is measurable. We also know that $|\mathcal{C}| = |\mathbb{R}|$, so $|\mathcal{P}(\mathcal{C})| = |\mathcal{P}(\mathbb{R})| = 2^{\mathbb{R}}$. Then

$$|\mathcal{M}(\mathbb{R})| \geq |\mathcal{P}(\mathcal{C})| = 2^{\mathbb{R}} = |\mathcal{P}(\mathbb{R})|,$$

showing that $|\mathcal{M}(\mathbb{R})| = |\mathcal{P}(\mathbb{R})|$.

**2.3.5. Regularity and Measurable Sets.** Let us now see that Lebesgue-measurable sets are nice.

***2.3.25. Proposition.***

*Let $E \subset \mathbb{R}$. The following statements are equivalent:*

*(i)* $E \in \mathcal{M}(\mathbb{R})$;

*(ii) for all* $\varepsilon > 0$, *there exists* $V \subset \mathbb{R}$, *open, such that* $E \subset V$, $m^*(V \setminus E) < \varepsilon$;

*(iii) for all* $\varepsilon > 0$, *there exists* $F \subset \mathbb{R}$, *closed, such that* $E \supset F$, $m^*(E \setminus F) < \varepsilon$.

*Proof.* Because (ii) and (iii) are related via taking complements, it seems necessary to pass through (i). Note that we use $m^*$ in the statements because when we go from (ii) and (iii) to (i) we don't know a priori that the sets we are dealing with are measurable.

(i) $\Longrightarrow$ (ii) Assume first that $m(E) < \infty$. By definition of the Lebesgue outer measure, there exists a countable open cover $\mathcal{O} = \{I_k\}$, such that $\ell(\mathcal{O}) - \varepsilon < m^*(E)$. Take $V = \bigcup_k I_k$. Then $V$ is open, $E \subset V$, and

$$m^*(V \setminus E) = m(V \setminus E) = m(V) - m(E) \leq \ell(\mathcal{O}) - m^*(E) < \varepsilon.$$

If $m(E) = \infty$, we can write $E = \bigcup_{k \in \mathbb{Z}} E_k$, where $E_k = E \cap [k, k+1)$. Using the first part of the proof, for each $k$ there exists $V_k$ open, with $E_k \subset V_k$, and $m(V_k \setminus E_k) < \frac{\varepsilon}{2^{|k|+2}}$. Put $V = \bigcup_k V_k$. Then $V$ is open, $E \subset V$ and

$$m^*(V \setminus E) \leq m^*\Big(\bigcup_k (V_k \setminus E_k)\Big) \leq \sum_k m^*(V_k \setminus E_k) \leq \sum_k \frac{\varepsilon}{2^{|k|+2}} < \varepsilon.$$

(ii) $\implies$ (i) Fix $S \subset \mathbb{R}$. Fix $\varepsilon > 0$. By hypothesis there exists $V$ open with $E \subset V$ and $m^*(V \setminus E) < \varepsilon$. Then

$$\begin{aligned} m^*(S) &\leq m^*(S \cap E^c) + m^*(S \cap E) \\ &= m^*(S \cap E^c \cap V) + m^*(S \cap E^c \cap V^c) + m^*(S \cap E) \\ &\leq m^*(E^c \cap V) + m^*(S \cap V^c) + m^*(S \cap V) \\ &= m^*(V \setminus E) + m^*(S) \leq \varepsilon + m^*(S). \end{aligned}$$

As this can be done for any $\varepsilon > 0$, we get that $m^*(S) = m^*(S \cap E^c) + m^*(S \cap E)$. Doing this for any $S \subset \mathbb{R}$, we get that $E \in \mathcal{M}(\mathbb{R})$.

(i) $\implies$ (iii) if $E \in \mathcal{M}(\mathbb{R})$, then so is $E^c$. Given $\varepsilon > 0$, we apply (ii) to obtain $V$ open, $E^c \subset V$, and $m(V \setminus E^c) < \varepsilon$. Now take $F = V^c$, closed, and note that $E \setminus F = V \setminus E^c$.

(iii) $\implies$ (i) Fix $\varepsilon > 0$. Taking $V = F^c$ we obtain that $E^c \subset V$ and $m(V \setminus E^c) < \varepsilon$. As this can be done for all $\varepsilon > 0$, we are in the situation of (ii) for the set $E^c$, and so by the above we conclude that $E^c \in \mathcal{M}(\mathbb{R})$. Thus $E \in \mathcal{M}(\mathbb{R})$. ☐

The above result is surprising, in the sense that it looks so similar to the definition of the Lebesgue outer measure $m^*$, that it almost looks like a proof that every set is measurable. After all, $m^*$ is defined by taking the infimum over open sets that contain $E$; so it is true that for any $E \subset \mathbb{R}$ there exists $V$ open with $m(V) - m^*(E) < \varepsilon$. The subtlety is in the fact that since $m^*$ is not a measure everywhere, said inequality tells us nothing about $m^*(V \setminus E)$. In fact, consider the Vitali set $E \subset [0,1]$ from Remark 2.3.18 with $m^*(E) = 1$. If $V$ is open and $V \supset E$, then $V \setminus E$ still contains all the sets $E_r$, which means that $m^*(V \setminus E) \geq 1$, in spite of the fact that $E \subset V$, $m^*(E) = 1$, and given any $\delta > 0$ we can choose $V$ with $m(V) < 1 + \delta$.

***2.3.26. Corollary. (Outer Regularity of Lebesgue Measure)***

*If $E \in \mathcal{M}(\mathbb{R})$, then*

$$m(E) = \inf\{m(V) : \ V \subset \mathbb{R}, \quad \textit{open}, \ E \subset V\}.$$

*Proof.* For any such $V$ we have $E \subset V$ and so $m(E) \leq m(V)$. Given $\varepsilon > 0$, by Proposition 2.3.25 there exists $V$ open with $E \subset V$ and $m(V) - m(E) < \varepsilon$. As this can be done for all $\varepsilon > 0$, we get that $m(E) = \inf\{m(V) : \ V \subset \mathbb{R}, \ \text{open}, E \subset V\}$. ☐

Corollary 2.3.26 can also be proven by using Proposition 1.8.1 and the definition of outer measure. In any case, the notion of outer regularity is a desirable property that is not valid for an arbitrary measure on $\mathcal{B}(\mathbb{R})$. For an easy example, the counting measure on $\mathbb{R}$ is not outer regular (Exercise 2.3.25). Regularity does not work for the Lebesgue measure if we switch the roles of "open" and "closed" (Exercise 2.3.21).

***2.3.27. Lemma.***

*Let $E \in \mathcal{M}(\mathbb{R})$. Then there exists $B \in \mathcal{B}(\mathbb{R})$ with $B \subset E$ and $m(B) = m(E)$.*

*Proof.* For each $k \in \mathbb{N}$, by Proposition 2.3.25 there exists $F_k$, closed, with $F_k \subset E$ and $m(E \setminus F_k) < \frac{1}{k}$. Let $B = \bigcup_{k=1}^{\infty} F_k$. Then $B \in \mathcal{B}(\mathbb{R})$ and $B \subset E$. We have

$$m(E) - m(B) = m(E \setminus B) = m(E \cap B^c) = m\Big(E \cap \bigcap_k F_k^c\Big)$$

$$= m\Big(\bigcap_k E \cap F_k^c\Big) \le m(E \setminus F_n) \le \frac{1}{n}.$$

This can be done for any $n$, so $m(B) = m(E)$. ☐

***2.3.28. Proposition.***

*Let $E \subset \mathbb{R}$. Then $E \in \mathcal{M}(\mathbb{R})$ if and only if there exist $B \in \mathcal{B}(\mathbb{R})$ and $E_0 \in \mathcal{M}(\mathbb{R})$, disjoint, with $E = B \cup E_0$ and $m(E_0) = 0$. The set $B$ can be chosen to be an $F_\sigma$ set (countable union of closed).*

*Proof.* If $E = B \cup E_0$ with $B \in \mathcal{B}(\mathbb{R})$ and $E_0 \in \mathcal{M}(\mathbb{R})$, then $E \in \mathcal{M}(\mathbb{R})$ as it is a $\sigma$-algebra and $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$.

Conversely, if $E \in \mathcal{M}(\mathbb{R})$ by Lemma 2.3.27 there exists $B \subset E$ with $B \in \mathcal{M}(\mathbb{R})$ and $m(B) = m(E)$. Put $E_0 = E \setminus B \in \mathcal{M}(\mathbb{R})$. Then $E = E_0 \cup B$, and $m(E_0) = m(E) - m(B) = 0$. ☐

The regularity properties shown above imply that the Lebesgue measure fits nicely with the topological structure of the real line. And this is a situation that appears often: we have a measure, and we want to do integration, over a set that also carries a topological structure. In light of that we make the following definitions. In most cases of interest, the $\sigma$-algebra $\mathcal{A}$ will be the Borel $\sigma$-algebra $\mathcal{B}(X)$.

**2.3.29. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space that also carries a locally compact Hausdorff topology with $\mathcal{B}(X) \subset \mathcal{A}$. The measure $\mu$ is said to be*

- ***Outer Regular**, if for $E \in \mathcal{A}$*
$$\mu(E) = \inf\{\mu(V) : \ E \subset V, \ V \text{ open}\};$$
- ***Inner Regular**, if for $E \in \mathcal{A}$*
$$\mu(E) = \sup\{\mu(K) : \ K \subset E, \ K \text{ compact}\};$$
- ***Radon** if it is outer and inner regular, and finite on compact sets.*

If a measure is inner regular, then it is finite on compact sets if and only if it is **locally finite**: for each $x \in X$ there exists $V$ open with $x \in V$ and $\mu(V) < \infty$ (Exercise 2.3.22).

*2.3.30. Examples.*

- The Lebesgue measure is a Radon measure. This is Proposition 2.3.25 together with the fact that any closed set in $\mathbb{R}^n$ is an increasing union of compact sets.
- On any Hausdorff topological space, the Dirac measure (Example 2.3.7) is a Radon measure (Exercise 2.3.23).
- If $X$ is a Polish space (separable, metric, compact, complete), then any finite measure on $\mathcal{B}(X)$ is Radon (Exercise 2.3.24).
- The counting measure on $\mathbb{R}^n$ is not Radon (Exercise 2.3.25).

**2.3.6. Exercises.**

(2.3.1) Let $X$ be a set. Show that $(X, \mathcal{P}(X), \mu)$ is a measure space, where $\mu$ is the counting measure, given by
$$\mu(S) = \begin{cases} |S|, & \text{if } S \text{ is finite} \\ \infty, & \text{if } S \text{ is infinite} \end{cases}$$
and $|S|$ denotes the number of elements in $S$.

(2.3.2) Let $X$ be a set and let
$$\mathcal{A} = \{A \subset X : \quad A \text{ is countable or } A^c \text{ is countable}\}.$$
(i) Show that $\mathcal{A}$ is a $\sigma$-algebra.

(ii) Show that $\mathcal{A}$ is the $\sigma$-algebra generated by the singletons (the family of all subsets of $A$ consisting of a single element).

(iii) Show that $\mathcal{A} = \mathcal{P}(X)$ if and only if $X$ is countable.

(2.3.3) Consider a set $X$ and $\mathcal{A}$ as in Exercise 2.3.2. Let

$$\mu(A) = \begin{cases} 0, & A \text{ countable} \\ 1, & \text{otherwise} \end{cases}$$

Show that $(X, \mathcal{A}, \mu)$ is a measure space.

(2.3.4) Let $\mathcal{A}$ be a $\sigma$-algebra. Show $|\mathcal{A}| \neq |\mathbb{N}|$. *(Hint: if $\mathcal{A}$ is countably infinite, consider for each $x \in X$ the smallest set in $\mathcal{A}$ that contains $x$)*

(2.3.5) Let $(X, \mathcal{A})$ be a measurable space. Define

$$\mu(A) = \begin{cases} 0, & A \ \text{finite} \\ \infty, & A \ \text{infinite} \end{cases}$$

Show that $\mu$ is always additive, and discuss when it is $\sigma$-additive.

(2.3.6) Complete the details in Example 2.3.7. That is, show that $\delta$ is a measure, and that $x_0$ is an atom.

(2.3.7) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $E \in \mathcal{A}$. Show that

$$\mathcal{A}_E = \{A \cap E : \ A \in \mathcal{A}\}$$

is a $\sigma$-algebra on $E$, and that $\mu_E(A) = \mu(A)$ defines a measure on $\mathcal{A}_E$.

(2.3.8) Let $(X, \mathcal{A}, \mu)$ be a measure space and $Y$ a set with $X \subset Y$. Show that there exists a measure space $(Y, \mathcal{A}', \mu')$ such that $\mu'(Y \setminus X) = 0$, $\mathcal{A} = \mathcal{A}'_X$, and $\mu = \mu'_X$.

(2.3.9) Let $\mathcal{M}$ be an infinite $\sigma$-algebra. Show that there exists nonempty $E \in \mathcal{M}$ such that $\mathcal{M}_{E^c}$ is infinite.

(2.3.10) Let $\mathcal{M}$ be an infinite $\sigma$-algebra. Show that $\mathcal{M}$ contains a pairwise disjoint sequence of sets. *(Hint: the naive approach does not work; instead, use Exercise 2.3.9)*

(2.3.11) Show that the equality

$$\mu\Big(\bigcap_k E_k\Big) = \lim_{k\to\infty} \mu(E_k)$$

can fail if $\{E_k\}$ is non-increasing but $\mu(E_k) = \infty$ for all $k$. Examples can be found in the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ with $\mu$ the counting measure.

(2.3.12) Let $X$ be a set and $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be given by

$$\mu^*(E) = \begin{cases} 0, & E = \varnothing \\ 1, & E \neq \varnothing \end{cases}$$

Show that $\mu^*$ is an outer measure and find $\mathcal{M}(X)$.

(2.3.13) Let $\nu^* : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ be given by

$$\nu^*(E) = \begin{cases} 0, & E = \varnothing \\ \frac{|E|}{1+|E|}, & E \ \text{finite} \\ 1, & E \ \text{infinite} \end{cases}$$

Show that $\nu^*$ is an outer measure and find $\mathcal{M}(X)$.

(2.3.14) Suppose that in Definition 2.3.11 we replace "countable cover" with "finite cover". Show that this would give $m^*(\mathbb{Q} \cap [0,1]) = 1$, and that this would imply that $m^*$ is not an outer measure.

(2.3.15) Let $E \subset \mathbb{R}$ such that $m^*(E) > 0$. Show that there exist $a, b \in E$ such that $a - b$ is irrational.

(2.3.16) Let $E \subset \mathbb{R}$ be the set of all numbers in $[0,1]$ that do not have a 1 anywhere in their decimal expansion. Is $E$ measurable? Find $m^*(E)$.

(2.3.17) Let $A, B \subset \mathbb{R}$ be Lebesgue measurable with $m(A) < \infty$. Show that the function $f : \mathbb{R} \to [0, \infty)$ given by $f(x) = m((A + x) \cap B)$ is continuous. *(Hint: the assertion is easier to prove for intervals)*

(2.3.18) Show that the relation $x \sim y$ if $x - y \in \mathbb{Q}$ is an equivalence relation in $\mathbb{R}$.

(2.3.19) Fix $c \in (0,1)$. Construct an open set $V \subset [0,1]$, dense in $[0,1]$ and with $m(V) = c$ *(Hint: in Exercise 2.2.8 this was done for $c = \frac{1}{2}$; an entirely different approach is possible with an idea similar—but not equal—to Exercise 2.3.17)* .

(2.3.20) Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$. Show that $\mu$ is outer regular if and only if it is inner regular by closed sets.

(2.3.21) Show that Lebesgue measure is not outer regular by closed sets, nor inner regular by open sets.

(2.3.22) Show that a Borel measure on a locally compact Hausdorff space is locally finite if and only if it is finite on compact sets.

(2.3.23) Show that if $X$ is a Hausdorff topological space, a Dirac delta is a Radon measure on $\mathcal{B}(X)$. Show by example that the assertion can fail if $X$ is not Hausdorff.

(2.3.24) Let $X$ be metric and compact. Show that any finite measure on $\mathcal{B}(X)$ is Radon.

(2.3.25) Show that the counting measure on $\mathbb{R}^n$ is not Radon.

(2.3.26) Let $(X, \mathcal{A}, \mu)$ be a measure space. The measure $\mu$ is **semifinite** if whenever $\mu(E) = \infty$ for some $E \in \mathcal{A}$, there exists $F \in \mathcal{A}$ with $F \subset E$ and $0 < \mu(F) < \infty$. Show that for such $\mu$ any $E \in \mathcal{A}$ satisfies

$$\mu(E) = \sup\{\mu(F) : \ F \in \mathcal{A}, \ F \subset E, \ \mu(F) < \infty\}. \tag{2.12}$$

Show also that the equality above fails for a measure that is not semifinite.

(2.3.27) Let $K$ be a topological space and $\{\mu_j\}$ a collection of Borel measures on $K$. Let $X = \bigcup_j K \times \{j\}$. We write $\pi_1$ for the coordinate function $\pi_1(a, b) = a$.

(i) Show that

$$\Sigma = \big\{B \subset X : \ \pi_1(B \cap (K \times \{j\})) \in \mathcal{B}(K) \ \text{ for all } j\big\}$$

is a $\sigma$-algebra.

(ii) Show that $\mu$ given by $\mu(B) = \sum_j \mu_j\big(\pi_1(B\cap(K\times\{j\}))\big)$ is a measure on $\Sigma$.

## 2.4. Measurable Functions

With the Lebesgue measure constructed and a bit of a basic theory of measure spaces, we can now move to the other part of the program outlined in Section 2.1. To be able to define an integral using (2.3), we need to discuss the functions that have measurable preimages of intervals.

**2.4.1. Definition.**

*Let $(X, \mathcal{A})$ be a measurable space, and $(Y, \mathcal{T})$ a topological space. A function $f : X \to Y$ is Borel-**measurable** if $f^{-1}(V) \in \mathcal{A}$ for all $V \in \mathcal{T}$.*

In the definition above, we have a measurable space in the domain of $f$, and a topological space in the codomain. This is natural, as explained in Section 2.1, because for integration the roles of the domain and codomain of the function are very different. That said, we will show in Proposition 2.4.3 that the definition can be rephrased in slightly more symmetric terms, and it fits what would be a general definition where a measurable function between two measurable spaces is precisely a function such that preimage of measurable is measurable. That is, one can define measurability for a function between two measurable spaces as a function such that the preimage of measurable is measurable, but we will not need such generality here. Here we focus on Borel measurability, which is equivalent to Definition 2.4.1 (see Proposition 2.4.3).

As Borel measurability is the only one we will consider here, we will drop the name Borel and just say that $f$ is measurable.

If $X$ has non-measurable subsets, non-measurable functions will exist. For we can take $E \subset X$ non-measurable, and consider $1_E : X \to \mathbb{R}$. Then $1_E^{-1}((0, 2)) = E$, which is not measurable.

Recall that if $(X, \mathcal{A}, \mu)$ is a measure space and $P(x)$ is a statement for each $x \in X$, we say that $P(x)$ is true *almost everywhere* (a.e.) if there exists $E \in \mathcal{A}$ such that $P(x)$ is true for all $x \in E$ and $\mu(E^c) = 0$. As has been mentioned, in measure theory things that happen almost everywhere can usually be considered as happening everywhere. For instance,

**2.4.2. Lemma.**

*If $f = g$ a.e. on a complete measure space and $g$ is measurable, then $f$ is measurable.*

*Proof.* Let $E = \{f = g\}$. Then $\mu(E^c) = 0$ (in particular, we are assuming that $E^c$ is measurable, which also makes $E$ measurable). For any open set $V$

$$\begin{aligned} f^{-1}(V) &= \{x \in E : \ f(x) \in V\} \cup \{x \in E^c : \ f(x) \in V\} \\ &= \{x \in E : \ g(x) \in V\} \cup \{x \in E^c : \ f(x) \in V\}. \end{aligned}$$

So $f^{-1}(V) = (E \cap g^{-1}(V)) \cup E'$, with $E$ measurable and $E' \subset E^c$ a nullset (hence measurable by completeness), making $f^{-1}(V)$ measurable. Then $f$ is measurable. ☐

We will need to test measurability both in bigger and smaller sets than the topology $\mathcal{T}$; for instance, in the real line we might want to test measurability on closed sets, or may just want to be able to test it on a family that generates the topology (like the open intervals).

**2.4.3. Proposition.**

*Let $(X, \mathcal{A})$ be a measurable space, and $(Y, \mathcal{T})$ a topological space. A function $f : X \to Y$ is measurable if and only if $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(Y)$.*

*Proof.* Since $\mathcal{T} \subset \mathcal{B}(Y)$, if $f^{-1}(B)$ is measurable for all Borel sets, it is in particular measurable for the open sets, and so $f$ is measurable.

Conversely, assume that $f$ is measurable and let $\mathcal{F} = \{E \subset Y : \ f^{-1}(E) \in \mathcal{A}\}$. Then $\mathcal{F}$ is a $\sigma$-algebra (Exercise 2.4.1). By hypothesis, $\mathcal{F}$ contains the open sets. Then $\mathcal{B}(Y) \subset \mathcal{F}$. ☐

**2.4.4. Proposition.**

*Let $(X, \mathcal{A})$ be a measurable space, and $(Y, \mathcal{T})$ a topological space. Let $\mathcal{E} \subset \mathcal{B}(Y)$ such that $\Sigma(\mathcal{E}) = \mathcal{B}(Y)$. A function $f : X \to Y$ is measurable if and only if $f^{-1}(B) \in \mathcal{A}$ for all $E \in \mathcal{E}$.*

*Proof.* If $f$ is measurable, then $f^{-1}(E) \in \mathcal{A}$ for all $E \in \mathcal{E}$ by Proposition 2.4.3.

Conversely, suppose that $f^{-1}(E) \in \mathcal{A}$ for all $E \in \mathcal{E}$. Using $\mathcal{F}$ as in the proof of Proposition 2.4.3, we have that $\mathcal{E} \subset \mathcal{F}$. Then $\mathcal{B}(Y) \subset \mathcal{F}$ and so $f$ is measurable. ☐

It is still not clear what (or whether) functions are measurable. But that will be changing very soon! For real-valued functions, we can use Proposition 2.4.4 to make measurability more concrete.

**2.4.5. Corollary.**

*Let $(X, \mathcal{A})$ be a measurable space, and $f : X \to \mathbb{R}$. The following statements are equivalent:*

*(i)* *$f$ is measurable;*

*(ii)* *for all $a, b \in \mathbb{R}$ with $a < b$, $f^{-1}(a, b)$ is measurable;*

*(iii)* *for all $a, b \in \mathbb{R}$ with $a < b$, $f^{-1}[a, b]$ is measurable;*

*(iv)* *for all $a \in \mathbb{R}$, $f^{-1}(a, \infty)$ is measurable;*

*(v)* *for all $a \in \mathbb{R}$, $f^{-1}(-\infty, a)$ is measurable;*

*(vi)* *for all $a \in \mathbb{Q}$, $f^{-1}(a, \infty)$ is measurable.*

*Proof.* In each case, we have a set $\mathcal{E}$ with $\Sigma(\mathcal{E}) = \mathcal{B}(\mathbb{R})$, so Proposition 2.4.4 applies. □

When we consider a topological space $X$ and its Borel $\sigma$-algebra, we get by definition that continuous functions are measurable. More generally,

**2.4.6. Proposition.**

*Let $(X, \mathcal{A})$ be a measurable space and $Y, Z$ topological spaces. Let $f : X \to Y$ be measurable, and $g : Y \to Z$ be continuous. Then $g \circ f$ is measurable.*

*Proof.* For any $V \subset Z$ open, the continuity of $g$ gives us $g^{-1}(V)$ open in $Y$. Then $(g \circ f)^{-1}(V) = f^{-1}(g^{-1}(V)) \in \mathcal{A}$. □

The simple argument in Proposition 2.4.6 actually shows that compositions of the form $g \circ f$ with $g$ Borel and $f$ measurable are measurable.

We cannot expect in general that the composition of measurable functions will be measurable. We even need to be careful what we mean here, since the way we defined measurability we require a topological space in the codomain, so we would be constrained to use the Borel $\sigma$-algebra in the middle space. Here is the canonical example, where we get a composition of the form $g \circ f$ with $g$ measurable and $f$ continuous, which is not measurable. Let $\gamma : [0, 1] \to [0, 2]$ be the function from the proof of Proposition 2.3.24, and let $\rho : [0, 2] \to [0, 1]$ be its inverse. We proved in Proposition 2.3.24 that $m(\gamma(\mathcal{C})) = 1$, and by

Lemma 2.3.23 there exists $V \subset \gamma(\mathcal{C})$, not measurable. Let $B = \gamma^{-1}(V) \subset \mathcal{C}$. We have $B \in \mathcal{M}(\mathbb{R})$ by Lemma 2.3.20; so $1_B$ is Lebesgue measurable. We have

$$(1_B \circ \rho)^{-1}(\{1\}) = \rho^{-1}(1_B^{-1}(\{1\}) = \gamma(B) = V,$$

not measurable. We are considering $X = [0,2]$, $Y = [0,1]$, $Z = \mathbb{R}$, with the Lebesgue $\sigma$-algebra in $X$ and $Y$.

**2.4.7. Corollary.**

*Let $(X, \mathcal{A})$ be a measurable space. Let $f, g : X \to \mathbb{R}$ be measurable, and $\lambda \in \mathbb{R}$. Then the functions $f + \lambda g$, $fg$, and $|f|$ are measurable. If $f \neq 0$ a.e., then $1/f$ is measurable.*

*Proof.* Consider the function $h : X \to \mathbb{R}^2$ that maps $x \longmapsto (f(x), g(x))$. We claim that this $h$ is measurable. Indeed, for any rectangle $(a, b) \times (c, d)$ we have

$$h^{-1}\big((a, b) \times (c, d)\big) = f^{-1}(a, b) \cap g^{-1}(c, d) \in \mathcal{A}.$$

As every open ball is a countable union of rectangles, the set of rectangles generates $\mathcal{B}(\mathbb{R}^2)$ and Proposition 2.4.4 applies to conclude that $h$ is measurable. Consider next the function $+ : \mathbb{R}^2 \to \mathbb{R}$, with $+(a, b) = a + b$. This function is continuous, and so measurable when we consider $\mathcal{B}(\mathbb{R}^2)$ as the $\sigma$-algebra in the domain. Then, by Proposition 2.4.6,

$$f + g = + \circ h$$

is measurable. Similarly, $fg = \times \circ h$ is measurable, and so are $|f| = |\cdot| \circ f$ and $\lambda g = M_\lambda \circ g$ (where $M_\lambda$ is the continuous function of multiplication by $\lambda$).

Finally, the function $1/f$. It might be the case the $1/f$ takes the value $\pm\infty$, but that does not affect measurability:

$$\begin{aligned}
\left(\frac{1}{f}\right)^{-1}(a, \infty) &= \left\{t \in \mathbb{R} : \; f(t) \neq 0, \; \frac{1}{f(t)} > a\right\} \\
&= \begin{cases} \left\{f < \frac{1}{a}\right\} \cap \{f > 0\}, & \text{if } a > 0 \\ \{f > 0\}, & \text{if } a = 0 \\ \{f > 0\} \cup \left[\left\{f < \frac{1}{a}\right\} \cap \{f < 0\}\right], & \text{if } a < 0 \end{cases}
\end{aligned}$$

In all three cases we get measurable sets, and Corollary 2.4.5 applies. ☐

So in the case where our measure space is $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$, we automatically get that all continuous functions are measurable, and from Corollary 2.4.7 all compositions of the form “rational function with continuous” are measurable. But there is more.

### *2.4.8. Corollary.*

*Let $(X, \mathcal{A})$ be a measurable space and $f, g : X \to \mathbb{R}$ be measurable. Then the functions $\max\{f, g\}$ and $\min\{f, g\}$ are measurable. In particular the positive and negative parts, $f^+ = \max\{f, 0\}$ and $f^- = -\min\{f, 0\}$ are measurable.*

*Proof.* We have

$$\max\{f, g\} = \frac{f + g + |f - g|}{2}, \qquad \min\{f, g\} = \frac{f + g - |f - g|}{2}. \qquad \square$$

From now on, we will start considering functions $f : X \to \overline{\mathbb{R}}$, the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. Naively one would think that we would need to redo all our measurability results for such functions. But, given $f : X \to \overline{\mathbb{R}}$, we can take $A = \{x \in X : |f(x)| = \infty\}$, and write

$$f = f\,1_{X \setminus A} + f\,1_A.$$

The first summand is a usual function $X \to \mathbb{R}$, and so all that matters is that the set $A$ is measurable. In summary, as long as $f$ takes the values $\pm\infty$ on a measurable set (equivalently, $f$ is finite on a measurable set), we can treat these functions normally. Most of the time $A$ will be nullset, which makes things even easier.

Let us keep expanding the family of measurable functions.

### *2.4.9. Proposition.*

*Let $\{f_n\}$ be such that $f_n : X \to \overline{\mathbb{R}}$ is measurable for all $n$. Then the functions $g_j : X \to \overline{\mathbb{R}}$, $j = 1, \ldots, 4$, where*

$$g_1(x) = \sup_n f_n(x), \quad g_2(x) = \inf_n f_n(x),$$
$$g_3(x) = \limsup_n f_n(x), \quad g_4(x) = \liminf_n f_n(x),$$

*are measurable.*

*Proof.* Given $a \in \mathbb{R}$,

$$g_1^{-1}(a, \infty) = \{x : \; g_1(x) > a\} = \bigcup_n \{x : \; f_n(x) > a\} = \bigcup_n f_n^{-1}(a, \infty)$$

is measurable. So $g_1$ is measurable. Now $g_2(x) = -\sup_n(-f_n)$ is measurable. Also, $h_m(x) = \sup\{f_n(x) : \; m \leq n\}$ is measurable, and then $g_3(x) = \inf_m h_m(x)$ is measurable. Finally, $g_4(x) = -\limsup_n[-f_n(x)]$ is measurable. $\square$

***2.4.10. Corollary.***

*Let $(X, \mathcal{A}, \mu)$ be a complete measure space, and $\{f_n\}$ be a sequence of measurable functions such that $f_n \to f$ pointwise a.e. Then $f$ is measurable.*

*Proof.* Let $E$ be the set where the sequence does not converge. By hypothesis, $E$ is nullset. As $\mu$ is complete, $E$ is measurable and so is $E^c$. We may replace each $f_n$ with $f_n\, 1_{E^c}$, which is measurable and agrees with $f_n$ on $E^c$ and is zero on $E$. The new sequence now converges everywhere. By Proposition 2.4.9,

$$\lim_n f_n(x) = \limsup_n f_n(x)$$

for all $x$, and so the limit is measurable. This limit agrees with $f$ a.e., so $f$ is measurable by Lemma 2.4.2. □

**2.4.1. Simple Functions.** When one builds Calculus, the "basic" functions are the polynomials, as these are the few functions over $\mathbb{R}$ that one can actually evaluate. And the "good" functions are the continuous ones, better if they are differentiable. That will change here. For starters, Lebesgue integration does not reward continuity the same way that Riemann integration does. But, more importantly, Lebesgue integration requires a measure space, so the domain of our functions may not consist of numbers—making polynomials meaningless—and it may not carry a topology—making continuous functions also meaningless.

Here are our new "good guys":

**2.4.11. Definition.**

*Let $(X, \mathcal{A})$ be a measurable space. A function $f : X \to \overline{\mathbb{R}}$ is* ***simple*** *if it is measurable and it takes finitely many values.*

Let us work out what a simple function looks like. If $f$ is a simple function on $X$, then $f(X) = \{\lambda_1, \ldots, \lambda_m\}$, with $\lambda_1, \ldots, \lambda_m \in \overline{\mathbb{R}}$ and distinct. As $f$ is measurable, the sets $E_n = f^{-1}(\{\lambda_n\})$ are measurable. They are also disjoint and cover all of $X$. Thus we may write

$$f = \sum_{n=1}^{m} \lambda_n\, 1_{E_n},$$

which is the way we usually look at simple functions. The convention is that $\lambda_1, \ldots, \lambda_m$ are distinct, although in some cases allowing repetitions is useful.

**2.4.12. Lemma.**

*Let $(X, \mathcal{A})$ be a measurable space, and $f, g : X \to \overline{\mathbb{R}}$ be simple. Then $f + g$, $fg$, and $|f|$ are simple. More generally, $\varphi \circ f$ is simple for any function $\varphi : \mathbb{R} \to \mathbb{R}$.*

*Proof.* In all cases bar the last one, it is evident that the resulting functions are measurable and they take finitely many values, so they are simple. The composition $\varphi \circ f$ takes its values in the same sets as $f$, so it is measurable. That is,

$$\varphi \circ f = \sum_{n=1}^{m} \varphi(\lambda_n)\, 1_{E_n}$$

is simple. □

Let us work out a more explicit proof of Lemma 2.4.12, as the ideas will be useful. Suppose that $f = \sum_{n=1}^{m} \lambda_n\, 1_{E_n}$ and $g = \sum_{k=1}^{r} \gamma_k\, 1_{F_k}$. To work with both functions together, the crucial observation is that we can write

$$f = \sum_{n=1}^{m} \sum_{k=1}^{r} \lambda_n 1_{E_n \cap F_k}, \qquad g = \sum_{n=1}^{m} \sum_{k=1}^{r} \gamma_k 1_{E_n \cap F_k}.$$

The equalities are guaranteed by the fact that $\bigcup_n E_n = \bigcup_k F_k = X$ and that the measurable sets $\{E_n \cap F_k\}_{n,k}$ are pairwise disjoint. Now it is clear that

$$f + g = \sum_{n=1}^{m} \sum_{k=1}^{r} (\lambda_n + \gamma_k)\, 1_{E_n \cap F_k}, \qquad fg = \sum_{n=1}^{m} \sum_{k=1}^{r} (\lambda_n \gamma_k)\, 1_{E_n \cap F_k}.$$

are simple functions.

The usefulness of simple functions comes from the following essential result:

**2.4.13. Theorem.**

*Let $(X, \mathcal{A})$ be a measurable space, and $f : X \to \overline{\mathbb{R}}$ measurable. Then there exists a sequence $\{s_n\}$ of simple functions $s_n : X \to \mathbb{R}$ with $s_n \to f$ pointwise. If $f \geq 0$, the sequence can be chosen to be non-decreasing. If $f$ is bounded, the sequence can be chosen so that the convergence is uniform.*

*Proof.* Assume first that $f \geq 0$. Define $s_n : X \to \mathbb{R}$ by

$$s_n(x) = \begin{cases} n, & \text{if } f(x) \geq n \\ \frac{k-1}{2^n}, & \text{if } f(x) \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right),\ k \in \{1, \ldots, n2^n\} \end{cases}$$

Equivalently,

$$s_n = n\,1_{f^{-1}[n,\infty]} + \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\,1_{f^{-1}\left[\frac{k-1}{2^n},\frac{k}{2^n}\right)}.$$

The functions $s_n$ are simple: it is clear that they take finitely many values, and they are measurable because $f$ is. It is not hard to see (Exercise 2.4.8) that, for those $x \in X$ with $f(x) \leq n$,

$$s_n(x) = \frac{\lfloor 2^n f(x) \rfloor}{2^n}.$$

Note that $s_{n+1}(x) \geq s_n(x)$ for all $x$. Indeed, if $f(x) \geq n+1$ then $s_{n+1}(x) = n+1 \geq n = s_n(x)$. If $n \leq f(x) < n+1$, then

$$s_{n+1}(x) = \frac{\lfloor 2^{n+1} f(x) \rfloor}{2^{n+1}} \geq \frac{\lfloor 2^{n+1} n \rfloor}{2^{n+1}} = n = s_n(x).$$

And if $f(x) < n$, choose $k \in \{1, \ldots, n2^n\}$ such that

$$\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}.$$

Then

$$\frac{2k-2}{2^{n+1}} \leq f(x) < \frac{2k}{2^{n+1}}.$$

So $s_{n+1}(x)$ is either $\frac{2k-2}{2^{n+1}}$ or $\frac{2k-1}{2^{n+1}}$, while $s_n(x) = \frac{k-1}{2^n}$. In either case, $s_n(x) \leq s_{n+1}(x)$.

Given $x \in X$, if $f(x) \in \mathbb{R}$ then for $n$ big enough we will always have $f(x) < n$. By definition of $s_n$,

$$|f(x) - s_n(x)| < \frac{1}{2^n}. \tag{2.13}$$

So $s_n(x) \to f(x)$. If $f(x) = \infty$, then $s_n(x) = n$ for all $n$ and so $s_n(x) \to f(x)$, giving us the pointwise convergence. When $f$ is bounded, (2.13) will hold for all $x$ if $n$ is big enough, showing that the convergence is uniform.

Now for general $f$, we can write $f = f^+ - f^-$, with $f^+, f^- \geq 0$. For each of these we can use the above to find sequences of simple functions $\{s_n^+\}$ and $\{s_n^-\}$ that converge pointwise (and uniformly if the corresponding function is bounded) to $f^+$ and $f^-$. Then $f = \lim_n s_n^+ - s_n^-$ pointwise, and uniformly if $f$ is bounded. □

The power of Theorem 2.4.13—in the case of the real line—lies in ditching intervals for more general sets. For non-continuous functions $f$, the sets that appear in the simple functions that approximate $f$ will be often far from being intervals. For an extreme case of this, consider the simple function $1_{\mathbb{Q}}$. From the naive calculus point of view, this function is discontinuous everywhere, so very pathological. From our new point of view, it is a simple function that takes only two values, so it is among the simplest of simple functions. It is common to distinguish simple functions from **step functions**. These are the simple functions which are linear combinations of characteristic functions of intervals—in the real line, of course, where intervals make sense. It can be shown that a uniform limit of step functions is Riemann-integrable, so using step functions instead of the more general simple functions is fairly restrictive.

**2.4.14. Definition.**

*On a measure space $(X, \mathcal{A}, \mu)$, the* ***essential range*** *of a measurable function $f : X \to \mathbb{C}$ is the set*

$$\operatorname{ess\,ran} f = \{\alpha \in \mathbb{C} : \text{ for all } \varepsilon > 0,\ \mu(f^{-1}(B_\varepsilon(\alpha))) > 0\}.$$

The essential range of a measurable function is always closed (Exercise 2.4.9). The proof of Theorem 2.4.13 can be slightly tweaked so that each $s_n$ takes values in $\operatorname{ess\,ran} f$.

**2.4.2. Exercises.**

(2.4.1) Let $(X, \mathcal{A})$ be a measurable spaces and $f : X \to Y$ a function. Show that
$$\mathcal{F} = \{E \subset Y : \ f^{-1}(E) \in \mathcal{A}\}$$
is a $\sigma$-algebra.

(2.4.2) Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and $f : X \to Y$ a function. Show that
$$\mathcal{A}_0 = \{f^{-1}(B) : \ B \in \mathcal{B}\}$$
is a $\sigma$-algebra. What is the relation between $\mathcal{A}_0$ and $\mathcal{A}$?

(2.4.3) Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing. Show that $f$ is Borel-measurable (that is, pre-image of open is Borel).

(2.4.4) Let $f : \mathbb{R} \to \mathbb{R}$ be non-decreasing. Show that $f$ is continuous almost everywhere, by showing that the set of discontinuity points of $f$ is at most countable.

(2.4.5) Let $f : \mathbb{R} \to \mathbb{C}$ be Lebesgue-measurable. Show that there exists $g : \mathbb{R} \to \mathbb{C}$, Borel-measurable, with $f = g$ a.e. *(Hint: do it first for $f \geq 0$)*

(2.4.6) Let $(X, \mathcal{A})$ be a measurable space. If $f : E \to \mathbb{R}$ is a measurable function relative to the measurable space $(E, \mathcal{A}_E)$ (cfr. Exercise 2.3.7), show that the extension $\tilde{f} : X \to \mathbb{R}$ of $f$ given by $\tilde{f} = f\,1_E$ is measurable with respect to $\mathcal{A}$.

(2.4.7) If $\tilde{f} : X \to \mathbb{R}$ is a measurable function with respect to $\mathcal{A}$, and if $f = \tilde{f}|_E$ (the restriction of $\tilde{f}$ to $E$), show that $f : E \to \mathbb{R}$ is measurable with respect $\mathcal{A}_E$.

(2.4.8) Let $(X, \mathcal{A})$ be a measurable space and $f : X \to [0, \infty]$ measurable. Show that if $s_n$ is as in Theorem 2.4.13, then for $x \in f^{-1}[0, n)$ we have
$$s_n(x) = \frac{\lfloor 2^n f(x) \rfloor}{2^n}.$$

(2.4.9) Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{C}$ measurable. Show that $\operatorname{ess\,ran} f$ is closed.

(2.4.10) Show an example of a measure space, $g$ measurable, and $f = g$ a.e. with $f$ not measurable.

(2.4.11) Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) = 1$. Show that the following statements are equivalent:

(i) $\mu(E) \in \{0, 1\}$ for all $E \in \mathcal{A}$;

(ii) if $f : X \to \mathbb{R}$ is measurable, there exists $c \in \mathbb{R}$ with $f = c$ a.e.

## 2.5. The Lebesgue Integral

We finally have all the necessary ingredients to define the Lebesgue integral. As discussed in Section 2.1 an integral should be a limit of sums of the form "value of the function times size of the region". That makes it obvious what the integral of a simple function should be.

### 2.5.1. The Lebesgue Integral.

**2.5.1. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, and $s : X \to \mathbb{R}$ a simple non-negative function, $s = \sum_{j=1}^n \lambda_j \, 1_{E_j}$. The* ***Lebesgue integral*** *of $s$ with respect to $\mu$ is*

$$\int_X s \, d\mu = \sum_{j=1}^n \lambda_j \, \mu(E_j).$$

*For measurable $f : X \to [0, \infty]$, its Lebesgue integral with respect to $\mu$ is*

$$\int_X f \, d\mu = \sup \Big\{ \int_X s \, d\mu : \ s \text{ simple}, 0 \leq s \leq f \Big\}.$$

*If $E \subset X$ is measurable, we define*

$$\int_E f \, d\mu = \int_X f \, 1_E \, d\mu.$$

*If $f : X \to \overline{\mathbb{R}}$ satisfies that either $\int_X f^+ \, d\mu < \infty$ or $\int_X f^- \, d\mu < \infty$, then we define*

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$

*For $f : X \to \mathbb{C}$ measurable, we define*

$$\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu$$

*if both integrals exist and are finite.*

For nonnegative measurable $f$, its Lebesgue integral always exists, whether finite or infinite. When $f$ takes positive and negative values (or non-real complex values) the existence of its Lebesgue integral depends on the integral of some or all of its positive/negative/real/imaginary parts being finite. The reason being that we cannot allow expressions of the form $\infty - \infty$ nor $r + i\infty$ nor $\infty + ir$ nor $\infty + i\infty$.

When necessary to emphasize the variable, we may write the Lebesgue integral as

$$\int_E f(x)\,d\mu(x).$$

And in cases where the context is clear, it is not uncommon to write

$$\int_X f, \qquad\qquad \text{or even} \qquad\qquad \int f.$$

It will require some work to unravel the definition of Lebesgue integral and to show that everything works as it should. Particularly, it is not obvious at all from the definition that the integral is **additive**, that is that $\int_X (f+g) = \int_X f + \int_X g$.

When the measure space is $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ and $f : \mathbb{R} \to \mathbb{R}$ is piecewise continuous with jump discontinuities and $f(x) = 0$ outside of $[a, b]$, with a bit more machinery (Exercise 2.5.11) it will be easy to check that the above definition gives

$$\int_{[a,b]} f\,dm = \int_a^b f(t)\,dt.$$

At the other end of the spectrum, let us see how different can a Lebesgue integral look from a Riemann integral. When $\mu$ is the counting measure on $\mathbb{N}$, for a nonnegative sequence $\{a_n\}$ we have (Exercise 2.5.1)

$$\int_{\mathbb{N}} a\,d\mu = \sum_{n=1}^{\infty} a_n.$$

So we see that choosing different measures can lead to very different-looking Lebesgue integrals, and this provides a level of generality that goes way beyond extending Riemann integration to a bigger family of functions. Eventually, one realizes that integration is anything but assigning a number to a function in a linear way; that is, integrals are *linear functionals*. This point of view will be explored with care much later in the text (among other places, in Sections 4.5 and 5.5, and most of the text beyond that; linear functionals are a regular staple in functional analysis).

For yet another example both of a Lebesgue integral and of the "linear functional" point of view, consider the map $\varphi$ on $C[0,1]$ that to any $f$ assigns $f(0)$. That is, $\varphi(f) = f(0)$. If $\delta$ denotes the Dirac measure on $[0,1]$ (Example 2.3.7), one can check (Exercise 2.5.5) that

$$f(0) = \varphi(f) = \int_{[0,1]} f\, d\delta. \tag{2.14}$$

Let us start by showing the most basic properties (but note that additivity is missing!).

**2.5.2. Proposition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $f, g : X \to \overline{\mathbb{R}}$ measurable, $A, B \in \mathcal{A}$.*

*(i) If $0 \leq f \leq g$, then $\int_X f\, d\mu \leq \int_X g\, d\mu$.*

*(ii) If $\lambda \in [0, \infty)$ and $f \geq 0$, then $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$.*

*(iii) If $A \subset B$ and $0 \leq f$, then $\int_A f\, d\mu \leq \int_B f\, d\mu$.*

*(iv) If $f = 0$ a.e. or $\mu(A) = 0$, then $\int_A f\, d\mu = 0$.*

*(v) If $f \geq 0$ a.e. and $\int_X f\, d\mu = 0$, then $f = 0$ a.e.*

*Proof.* All these properties follow by definition.

(i) Every simple function with $0 \leq s \leq f$ satisfies $0 \leq s \leq g$, so in the integral for $g$ we are taking the supremum of a larger set.

(ii) The equality is trivial when $f$ is simple, and for general $f \geq 0$ we just need to take the non-negative scalar $\lambda$ out of a supremum of non-negative numbers.

(iii) If $A \subset B$ and $f \geq 0$, then $f\,1_A \leq f\,1_B$, so the inequality follows from (i).

(iv) If $f = 0$ a.e., let $E = \{f = 0\}$, so that $\mu(E^c) = 0$. Let $s = \sum_{k=1}^m \lambda_k\, 1_{E_k}$ be a simple function with $0 \leq s \leq f$. Since $s = 0$ on $E$, we may assume that $\lambda_1 = 0$ and $E_1 \supset E$. Then $E_2 \cup \cdots E_m \subset E^c$, which guarantees that $\mu(E_k) = 0$ for all $k \geq 2$. Hence

$$\int_X s\, d\mu = 0 + \sum_{k=2}^m \lambda_k\, \mu(E_k) = 0 + 0 = 0.$$

Thus the integral of $f$ is the supremum of the set $\{0\}$; that is, $\int_X f\,d\mu = 0$.

For the second part, if $\mu(A) = 0$ then $f\,1_A = 0$ a.e., so

$$\int_A f\,d\mu = \int_X f\,1_A\,d\mu = 0.$$

(v) Exercise 2.5.3.

□

We still don't have additivity, but we are slowly moving towards it. First, for simple functions.

***2.5.3. Lemma.***

*The Lebesgue integral is additive on simple functions.*

*Proof.* Let $f, g : X \to \mathbb{R}$ be simple functions on the measure space $(X, \mathcal{A}, \mu)$,

$$f = \sum_{j=1}^{m} \alpha_j\,1_{E_j}, \qquad g = \sum_{k=1}^{n} \beta_k\,1_{F_k}.$$

Using that $\bigcup_j E_j = \bigcup_k F_k = X$, we can write

$$f = \sum_{j=1}^{m}\sum_{k=1}^{n} \alpha_j\,1_{E_j \cap F_k}, \qquad g = \sum_{j=1}^{m}\sum_{k=1}^{n} \beta_k\,1_{E_j \cap F_k}.$$

Then

$$\begin{aligned}
\int_X (f+g)\,d\mu &= \sum_{j=1}^{m}\sum_{k=1}^{n} (\alpha_j + \beta_k)\,\mu(E_j \cap F_k) \\
&= \sum_{j=1}^{m} \alpha_j \sum_{k=1}^{n} \mu(E_j \cap F_k) + \sum_{k=1}^{n} \beta_k \sum_{j=1}^{m} \mu(E_j \cap F_k) \\
&= \sum_{j=1}^{m} \alpha_j\,\mu(E_j) + \sum_{k=1}^{n} \beta_k\,\mu(F_k) \\
&= \int_X f\,d\mu + \int_X g\,d\mu.
\end{aligned}$$

□

The jump from additivity for simple functions to general functions is not trivial. It will require us to prove the first big convergence theorem.

**2.5.2. The Convergence Theorems.** A thing one learns early in analysis is to be scared of exchanging limits; and rightly so, as the easy example

$$\lim_{x\to 0}\lim_{y\to 0}\frac{x}{x+y} = 1 \neq 0 = \lim_{y\to 0}\lim_{x\to 0}\frac{x}{x+y}$$

shows. And both series and integrals are limits, so exchanging a limit with a series or an integral should be seen with suspicion. We will see now two fairly general results that allow us to do the exchange between limit and integral.

***2.5.4. Theorem. (Monotone Convergence)***

*Let $(X,\mathcal{A},\mu)$ be a measure space, $E\in\mathcal{A}$, and $f_n : X\to[0,\infty]$, measurable for all $n\in\mathbb{N}$ and such that $f_n(x)\leq f_{n+1}(x)$ almost everywhere on $E$. Then*

$$\lim_{n\to\infty}\int_E f_n\,d\mu = \int_E\big(\lim_{n\to\infty} f_n\big)\,d\mu \tag{2.15}$$

*Proof.* We may change $E$ up to a nullset and the $f_n$ so that the inequalities hold for all $x$, without altering the integrals. Concretely, if $E_n$ is the nullset $\{f_n > f_{n+1}\}$, the union $\bigcup_n E_n$ is a nullset where we can redefine $f_n|_{\bigcup_n E_n} = 0$.

Let $f=\lim_n f_n$. It exists—possibly being infinite—by the monotonicity of $\{f_n\}$. Let

$$\alpha = \lim_{n\to\infty}\int_E f_n\,d\mu.$$

The limit exists in $\overline{\mathbb{R}}$ because the integrals form a monotone sequence in $\overline{\mathbb{R}}$. Because $f_n\leq f$ for all $n$, we have $\alpha\leq\displaystyle\int_E f\,d\mu$. Then, if $\alpha=\infty$, we are done. So we assume $\alpha<\infty$. Fix $c\in(0,1)$ and $s:X\to\mathbb{R}$, simple, with $0\leq s\leq f$. Let

$$E_n = \{x\in E:\ f_n(x)\geq c\,s(x)\}.$$

The monotonicity of the sequence guarantees that $E_1\subset E_2\subset\cdots$, and that $E=\displaystyle\bigcup_n E_n$; because for $n$ big enough we will have $f_n(x)\geq c\,s(x)$, since $f(x) > c\,s(x)$ when $s(x)>0$. As $E_n\subset E$ for all $n$,

$$\int_{E_n} s\,d\mu \leq \int_{E_{n+1}} s\,d\mu \leq \int_E s\,d\mu, \qquad n\in\mathbb{N}.$$

Because $s$ is finite-valued, there exists $r>0$ with $s(x)\leq r$ for all $x$. Then

$$\begin{aligned}\int_E s\,d\mu - \int_{E_n} s\,d\mu &= \int_X s\,(1_E - 1_{E_n})\,d\mu = \int_X s\,1_{E\setminus E_n}\,d\mu = \int_{E\setminus E_n} s\,d\mu\\ &\leq \int_{E\setminus E_n} r\,d\mu = r\,\mu(E\setminus E_n)\end{aligned}$$

$$= r\left[\mu(E) - \mu(E_n)\right] \xrightarrow[n\to\infty]{} 0$$

by continuity of the measure. Thus

$$c \int_E s\, d\mu = \lim_n \int_{E_n} cs\, d\mu \leq \limsup_n \int_{E_n} f_n\, d\mu \leq \limsup_n \int_E f_n = \alpha.$$

As $c$ was arbitrary, we get that $\int_E s\, d\mu \leq \alpha$. And as we can do that for any simple function $s$ with $0 \leq s \leq f$, we have shown that $\int_E f \leq \alpha$. This was the reverse inequality we needed, so the theorem is proved. □

An easy example—that was mentioned on page 82—where monotone convergence fails for the Riemann integral is the following. Let $\{q_n\}$ be an enumeration of $\mathbb{Q} \cap [0,1]$. Let $f_n = 1_{\{q_1,\ldots,q_n\}}$. Then $f_n \nearrow 1_{\mathbb{Q}}$, but the integral of the limit does not exist. Of course, for Riemann integrable functions the Lebesgue integral agrees with the Riemann integral, so in situations where Riemann integrable functions converge monotonically to a Riemann integrable function, we can apply Monotone Convergence to Riemann integrals.

An essential first application of Monotone Convergence is to show that the Lebesgue integral is indeed additive, which we will prove after an easy corollary.

**2.5.5. Corollary.**

*Let $f : X \to \overline{\mathbb{R}}$ be measurable, and let $\{E_n\}$ be an increasing union of measurable sets with union $X$. Then*

$$\int_X f\, d\mu = \lim_n \int_{E_n} f\, d\mu$$

*if the integral on the left exists.*

*Proof.* If $f \geq 0$, then Monotone Convergence gives us

$$\lim_n \int_{E_n} f\, d\mu = \lim_n \int_X f\, 1_{E_n}\, d\mu = \int_X f\, \lim_n 1_{E_n}\, d\mu = \int_X f\, d\mu,$$

since $\lim_n 1_{E_n} = 1$. Now for arbitrary $f$ such that its integral exists,

$$\int_X f\, d\mu = \int_X f^+\, d\mu - \int_X f^-\, d\mu = \lim_n \int_{E_n} f^+\, d\mu - \int_{E_n} f^-\, d\mu = \lim_n \int_{E_n} f\, d\mu.$$

□

Finally we can show the additivity. Note that also multiplication by arbitrary scalars had to wait up to now.

**2.5.6. Corollary.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f, g : X \to \overline{\mathbb{R}}$ measurable. If the integrals of $f$ and $g$ are finite then the integral of $f + \lambda g$ exists for all $\lambda \in \mathbb{R}$, and*

$$\int_X (f + \lambda g)\, d\mu = \int_X f\, d\mu + \lambda \int_X g\, d\mu. \tag{2.16}$$

*When $f, g \geq 0$ a.e. and $\lambda \geq 0$ the equality (2.16) holds even if some or all the integrals infinite. In particular, if $A, B \in \mathcal{A}$ with $A \cap B = \varnothing$, then*

$$\int_{A\cup B} f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu. \tag{2.17}$$

*Proof.* Assume first that $f, g \geq 0$ and $\lambda = 1$. Using Theorem 2.4.13 we construct sequences $\{s_n^f\}$ and $\{s_n^g\}$ of simple functions with $0 \leq s_n^f \nearrow f$ and $0 \leq s_n^g \nearrow g$. Then, using Monotone Convergence in the first and last equalities,

$$\begin{aligned}\int_X (f+g)\, d\mu &= \lim_n \int_X (s_n^f + s_n^g)\, d\mu = \lim_n \int_X s_n^f\, d\mu + \int_X s_n^g\, d\mu \\ &= \lim_n \int_X s_n^f\, d\mu + \lim_n \int_X s_n^g\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu.\end{aligned}$$

Now for arbitrary $f, g$ such that both $f$ and $g$ have finite integral we have that $f+g$ has finite integral, for $(f+g)^+ \leq f^+ + g^+$ and $(f+g)^- \leq f^- + g^-$. Also,

$$(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-. \tag{2.18}$$

We may rewrite (2.18) as

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+.$$

All six functions in this last equality are non-negative and have finite integral. So integrating and applying the first part of the proof, we get

$$\int_X (f+g)^+\, d\mu + \int_X f^-\, d\mu + \int_X g^-\, d\mu = \int_X (f+g)^-\, d\mu + \int_X f^+\, d\mu + \int_X g^+\, d\mu.$$

As all six integrals are finite, reordering we obtain

$$\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu.$$

If $A \cap B = \varnothing$, then

$$\int_{A\cup B} f\, d\mu = \int_X f\, 1_{A\cup B}\, d\mu = \int_X (f\, 1_A + f\, 1_B)\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu.$$

When some of the integrals are infinite but $f(X) \cup g(X) \subset \mathbb{R}$, let $E_n = \{x : |f(x)| + |g(x)| \leq n\}$. These form an increasing family of measurable sets with union $X$. By the above and Corollary 2.5.5,

$$\int_X (f+g)\,d\mu = \lim_n \int_{E_n} (f+g)\,d\mu = \lim_n \int_{E_n} f\,d\mu + \int_{E_n} g\,d\mu$$
$$= \int_X f\,d\mu + \int_X g\,d\mu.$$

Finally, if $f$ or $g$ is $\infty$ on a set of positive measure, then both sides of (2.16) are infinity and so the equality holds. Same if $f$ or $g$ is $-\infty$ on a set of positive measure. The case where $f=\infty$ on a set of positive measure and $g=-\infty$ on a set of positive measure—or viceversa—is precluded by the requirement that the integral of $f+g$ exists.

The trick we used above can be slightly generalized as follows: if $f = \sum_{k=1}^m f_k - \sum_{j=1}^n g_j$, with $f_k \geq 0$ and $g_j \geq 0$ a.e. for all $k, j$, and if $\int_X f_k\,d\mu < \infty$ for all $k$ or $\int_X g_j\,d\mu < \infty$ for all $j$, then

$$\int_X f\,d\mu = \sum_{k=1}^m \int_X f_k\,d\mu - \sum_{j=1}^n \int_X g_j\,d\mu.$$

Now we can deal with arbitrary scalars. Namely, if $f$ is measurable and its integral exists, and $\lambda \in \mathbb{R}$, we may write $\lambda = \lambda^+ - \lambda^-$ with $\lambda^+, \lambda^- \geq 0$, and then

$$\begin{aligned}\int_X \lambda f\,d\mu &= \int_X \left(\lambda^+ f^+ + \lambda^- f^- - \lambda^+ f^- - \lambda^- f^+\right) d\mu \\ &= \lambda^+ \int_X f^+\,d\mu + \lambda^- \int_X f^-\,d\mu - \lambda^+ \int_X f^-\,d\mu - \lambda^- \int_X f^+\,d\mu \\ &= (\lambda^+ - \lambda^-) \int_X (f^+ - f^-)\,d\mu = \lambda \int_X f\,d\mu.\end{aligned}$$

□

When $f$ is complex-valued, we write $f = g + ih$ with $g, h$ real. We defined

$$\int_X f\,d\mu = \int_X g\,d\mu + i \int_X h\,d\mu \tag{2.19}$$

when both the integrals of $g$ and $h$ exist as numbers. For complex numbers, allowing $\pm\infty$ as we did with the real line does not work well, so we require all integrals to be finite. Because the real part of the integral of $f$ is—by definition—the integral of the real part of $f$, and the same for the imaginary part, everything we have done so far works the same for a complex-valued functions, as long as both integrals in (2.19) exist as numbers. We even get

$$\overline{\int_X f\,d\mu} = \int_X \overline{f}\,d\mu.$$

directly from (2.19). The only thing that "mixes" real and imaginary parts so far is multiplication by a scalar. But the same idea from the proof of Corollary 2.5.6

can be used to extend the equality $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$ to complex valued $f$ and $\lambda \in \mathbb{C}$. Indeed, we write $f = g^+ - g^- + ih^+ - ih^-$, and similarly $\lambda = a + ib = a^+ - a^- + ib^+ - ib^-$. This allows us to write

$$\begin{aligned}\lambda f &= a^+g^+ - a^+g^- - a^-g^+ + a^-g^- - b^+h^+ + b^+h^- + b^-h^+ - b^-h^- \\ &\qquad + i\big[a^+h^+ - a^+h^- - a^-h^+ + a^-h^- + b^+g^+ - b^+g^- - b^-g^+ + b^-g^-\big].\end{aligned}$$

Then, as in the proof of Corollary 2.5.6, we can take non-negative scalars out of the integrals with non-negative functions and reconstruct the scalar outside of the integral, to get $\int_X \lambda f\, d\mu = \lambda \int_X f\, d\mu$.

We will continue to mostly work with real-valued functions. As mentioned, most results and operations do not mix the real and imaginary parts, so the results extend automatically to the complex case. An exception to this, multiplication by scalars, was just mentioned, and a second exception is the absolute value; we will address it in Lemma 2.5.13.

*2.5.7. Example.* The Monotone Convergence Theorem has both theoretical and practical applications. Here is a simple example of a practical application. Given $p > 0$,

$$\begin{aligned}\int_0^1 \frac{x^{p-1}}{1-x} \log x\, dx &= -\int_0^1 \Big(\sum_{k=0}^{\infty} -x^{p-1+k} \log x\Big)\, dx \\ &= -\sum_{k=0}^{\infty} \int_0^1 -x^{p-1+k} \log x\, dx \\ &\text{(Monotone Convergence)} \\ &= -\sum_{k=0}^{\infty} \frac{1}{(k+p)^2}.\end{aligned}$$

The series has positive terms, so the sequence of partial sums is monotone non-decreasing and non-negative (due to the minus sign), so Monotone Convergence applies. The case $p = 1$ of the series is well-known (see Example 3.10.6 and/or Example 8.1.7), so we get

$$\int_0^1 \frac{\log x}{1-x}\, dx = -\frac{\pi^2}{6} \qquad \text{and} \qquad \sum_{k=n}^{\infty} \frac{1}{k^2} = -\int_0^1 \frac{x^{n-1}}{1-x} \log x\, dx.$$

Let us formally write the idea used in Example 2.5.7.

**2.5.8. Corollary.**

*If $(X, \mathcal{A}, \mu)$ is a measure space and $f_n : X \to [0, \infty]$ is measurable for all $n$, then*

$$\int_X \Big( \sum_{n=1}^{\infty} f_n \Big) \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu.$$

*Proof.* We apply Monotone Convergence to the sequence $\{g_k\}$, where $g_k = \sum_{n=1}^{k} f_n$:

$$\begin{aligned} \int_X \Big( \sum_{n=1}^{\infty} f_n \Big) \, d\mu &= \int_X \big( \lim_k g_k \big) \, d\mu = \lim_k \int_X g_k \, d\mu \\ &= \lim_k \sum_{n=1}^{k} \int_X f_n \, d\mu = \sum_{n=1}^{\infty} \int_X f_n \, d\mu. \end{aligned}$$ □

**2.5.9. Corollary.**

*Let $f : [a, b] \to \mathbb{R}$ be piecewise continuous with jump discontinuities (even infinitely many). Then*

$$\int_{[a,b]} f \, dm = \int_a^b f(t) \, dt.$$

*Proof.* Exercise 2.5.11. □

While every Riemann integrable function is Lebesgue integrable and the Riemann integral is a Lebesgue integral, some issues arise when considering improper integrals. The problem is that improper Riemann integrals allow conditional convergence, while the Lebesgue integral behaves like absolute convergence. For a concrete example of this, the improper Riemann integral

$$\int_0^{\infty} \frac{\sin x}{x} \, dx$$

exists, but it does not exist as a Lebesgue integral. This is because

$$\begin{aligned} \int_0^{\infty} \frac{\sin^+ x}{x} \, dx &= \sum_{k=0}^{\infty} \int_{2k\pi}^{(2k+1)\pi} \frac{\sin x}{x} \, dx \geq \sum_{k=0}^{\infty} \int_{2k\pi + \frac{\pi}{4}}^{(2k+1)\pi - \frac{\pi}{4}} \frac{\sin x}{x} \, dx \\ &\geq \frac{\sqrt{2}}{2} \sum_{k=0}^{\infty} \int_{2k\pi + \frac{\pi}{4}}^{(2k+1)\pi - \frac{\pi}{4}} \frac{1}{x} \, dx \geq \frac{\pi\sqrt{2}}{4} \sum_{k=0}^{\infty} \frac{1}{(2k+1)\pi - \frac{\pi}{4}} \end{aligned}$$

$$\geq \frac{\pi\sqrt{2}}{4} \sum_{k=0}^{\infty} \frac{1}{(2k+1)\pi} = \infty,$$

and similarly with the $\sin^- x$ part. Interestingly, further measure-theoretic results will allow us to calculate the value of this integral (Exercise 2.7.14).

*2.5.10. Example.* The "non-decreasing" hypothesis in Theorem 2.5.4 is not just a gimmick of the proof. Consider, on the real line, the functions $g_k = 1_{[k,\infty)}$. This is a decreasing sequence of non-negative functions. We have

$$\int_{\mathbb{R}} g_k \, dm = \infty \quad \text{for all } k. \qquad \text{while} \qquad \int_{\mathbb{R}} \lim_k g_k \, dm = \int_{\mathbb{R}} 0 \, dm = 0.$$

We continue with the next result about exchanging limits and integrals.

***2.5.11. Theorem. (Fatou's Lemma)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f_n : X \to [0, \infty]$ measurable for all $n$. Then*

$$\int_X \liminf_n f_n \, d\mu \leq \liminf_n \int_X f_n \, d\mu.$$

*Proof.* Let $g_n(x) = \inf\{f_k(x) : k \geq n\}$, $n \in \mathbb{N}$. Then $g_n$ is measurable (Proposition 2.4.9), and by definition of the limit inferior $\liminf_n f_n = \sup_n g_n = \lim_n g_n$. When $n \leq k$,

$$\int_X g_n \, d\mu \leq \int_X f_k \, d\mu, \qquad \text{so} \qquad \int_X g_n \, d\mu \leq \inf_{k \geq n} \int_X f_k \, d\mu.$$

Then, using that $\{g_n\}$ is non-negative and non-decreasing so Monotone Convergence applies,

$$\begin{aligned} \int_X \liminf_n f_n \, d\mu &= \int_X \sup_n g_n \, d\mu = \int_X \lim_n g_n \, d\mu = \lim_n \int_X g_n \, d\mu \\ &\leq \lim_n \inf_{k \geq n} \int_X f_k \, d\mu = \liminf_n \int_X f_n \, d\mu. \end{aligned}$$ $\square$

The inequality in Fatou's Lemma can be strict in any nontrivial measure space. Indeed, if the space $(X, \mathcal{A}, \mu)$ has a measurable set $E$ with $\mu(E) > 0$ and $\mu(E^c) > 0$ (even if one or both are infinite), define

$$f_n = \begin{cases} 1_E, & n \text{ odd} \\ 1_{E^c}, & n \text{ even} \end{cases}$$

Then $\liminf_n f_n = 0$, so $\int_X \liminf_n f_n \, d\mu = 0$, while

$$\liminf_n \int_X f_n \, d\mu = \min\{\mu(E), \mu(E^c)\} > 0.$$

**2.5.12. Definition.**

*A measurable function $f : X \to \overline{\mathbb{R}}$ is **integrable** if* $\displaystyle\int_X |f| \, d\mu < \infty$.

As $|f| = f^+ + f^-$, for $f$ to be integrable is the same as requiring that $\int_X f^+ \, d\mu < \infty$ and $\int_X f^- \, d\mu < \infty$. It follows from the triangle inequality that linear combinations of integrable functions are integrable.

We record here a couple of basic observations, to be used often:

***2.5.13. Lemma.***

*If $f$ is measurable, real or complex valued, and* $\displaystyle\int_X f \, d\mu$ *exists, then*

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu. \tag{2.20}$$

*If $f$ is integrable, then $|f| < \infty$ a.e.*

*Proof.* We do the real-valued case first. Since the integral of $f$ exists, the following estimations work even if one of the terms is $\pm\infty$.

$$\left| \int_X f \, d\mu \right| = \left| \int_X f^+ \, d\mu - \int_X f^- \, d\mu \right| \leq \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \int_X |f| \, d\mu.$$

When $f$ is complex-valued, we write $\displaystyle\left| \int_X f \, d\mu \right| = \lambda \int_X f \, d\mu$ an appropriate $\lambda \in \mathbb{T}$ (don't forget that an integral is nothing but a number!). Then

$$\left| \int_X f \, d\mu \right| = \lambda \int_X f \, d\mu = \int_X \lambda f \, d\mu = \operatorname{Re} \int_X \lambda f \, d\mu = \int_X \operatorname{Re}(\lambda f) \, d\mu$$
$$\leq \int_X |\operatorname{Re}(\lambda f)| \, d\mu \leq \int_X |\lambda f| \, d\mu = \int_X |f| \, d\mu.$$

The second assertion in the lemma is Exercise 2.5.3. $\square$

The inequality (2.20) should be recognized as a generalization of the triangle inequality.

***2.5.14. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space and $f_n : X \to [0, \infty]$ measurable and such that $f_n \leq f_{n+1}$ a.e. for all $n$. If $\sup_n \int_X f_n \, d\mu < \infty$, then $f = \lim_n f_n$ is finite almost everywhere, integrable, and*

$$\int_X f \, d\mu = \lim_n \int_X f_n \, d\mu.$$

*Proof.* Let $E_n = \{f_n \not\leq f_{n+1}\}$. By hypothesis, $\mu(E_n) = 0$ for all $n$. Then $\mu\big(\bigcup_n E_n\big) = 0$. If we re-define each $f_n$ to be zero on $\bigcup_n E_n$, now the sequence is monotone everywhere, and the integrals are still the same.

Let $f(x) = \sup_n f_n(x)$. The monotonicity of the sequence guarantees that the limit exists, even if infinite. Let $C > 0$ such that $\int_X f_n \, d\mu < C$ for all $n$. By Monotone Convergence,

$$\int_X f \, d\mu = \lim_n \int_X f_n \, d\mu < C.$$

By Lemma 2.5.13 we get that $f < \infty$ a.e. $\square$

Now the second big integration theorem.

***2.5.15. Theorem. (Dominated Convergence)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $E \in \mathcal{A}$, $\{f_n\}_{n\in\mathbb{N}}$ measurable functions with $f_n \to f$ a.e. If there exists $g : X \to [0, \infty]$ integrable with $|f_n| \leq g$ a.e. for all $n$, then*

$$\lim_n \int_E |f - f_n| \, d\mu = 0. \tag{2.21}$$

*In particular,*

$$\lim_n \int_E f_n \, d\mu = \int_E f \, d\mu. \tag{2.22}$$

*Proof.* Since we may replace $f_n$ with $f_n 1_E$ and $g$ with $g 1_E$, we may assume without loss of generality that $E = X$. From $|f_n| \leq g$ we have $|f| \leq g$ and that each $f_n$ is integrable, and so $f_n + g$ is integrable. As $|f_n - f| \leq 2g$, the function $2g - |f_n - f| \geq 0$ is integrable. Using Fatou's Lemma,

$$\begin{aligned}\int_X 2g \, d\mu &= \int_X \liminf_n (2g - |f - f_n|) \, d\mu \leq \liminf_n \int_X (2g - |f - f_n|) \, d\mu \\ &= \int_X 2g \, d\mu + \liminf_n \Big[ -\int_X |f - f_n| \, d\mu \Big]\end{aligned}$$

$$= \int_X 2g\,d\mu - \limsup_n \int_X |f - f_n|\,d\mu.$$

It follows that

$$0 \le \limsup_n \int_X |f - f_n|\,d\mu \le 0,$$

showing (2.21). For (2.22),

$$\left|\int_X f_n\,d\mu - \int_X f\,d\mu\right| \le \int_X |f_n - f|\,d\mu \xrightarrow[n\to\infty]{} 0. \qquad \square$$

It is important to note that both Fatou's Lemma and Dominated Convergence work for sequences, but they can fail for nets. For instance let $\mathcal{F} = \{F \subset [0,1] : \text{ finite}\}$, ordered by inclusion. The net $\{1_F\}_{F\in\mathcal{F}}$ is increasing (the order is not total, though), bounded by an integrable function (the constant function 1) and converges pointwise to 1. But we have

$$\lim_F \int_{[0,1]} 1_F\,dm = 0, \qquad \int_{[0,1]} \lim_F 1_F\,dm = 1.$$

These functions feel a bit like cheating, since $1_F = 0$ a.e. But we can tweak the example bit. Since each $F \in \mathcal{F}$ is finite, it is possible to choose a continuous function $f_F$, such that $f_F(x) = 1$ for all $x \in F$, $|f_F| \le 1$ for all $F$, and $\int_{[0,1]} f\,dm = \frac{1}{2}$. Then the net $\{f_F\}$ of continuous functions also fails Dominated Convergence, as

$$\int_{[0,1]} \lim_F f_F\,dm = 1 \ne \frac{1}{2} = \lim_F \int_{[0,1]} f_F\,dm.$$

*2.5.16. Example.* A typical application of Dominated Convergence. Consider

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \frac{n \sin\frac{x}{n}}{x(x^2+1)}\,dx.$$

We want to apply Dominated Convergence. Since $\frac{\sin t}{t}$ has a removable discontinuity at 0, we can think of it as a continuous function with limits zero at $\pm\infty$; and $\left|\frac{\sin t}{t}\right| \le 1$ for all $t$. Then

$$\frac{n \sin\frac{x}{n}}{x(x^2+1)} = \frac{\sin\frac{x}{n}}{\frac{x}{n}} \frac{1}{x^2+1} \le \frac{1}{x^2+1}.$$

As our integrand is bounded by the integrable function $g(x) = \frac{1}{x^2+1}$, Dominated Convergence applies and we have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \frac{n \sin\frac{x}{n}}{x(x^2+1)}\,dx = \int_{-\infty}^{\infty} \lim_{n\to\infty} \frac{n \sin\frac{x}{n}}{x(x^2+1)}\,dx = \int_{-\infty}^{\infty} \frac{1}{x^2+1}\,dx = \pi.$$

We finish this section with the following result from Lebesgue. The key idea for the proof is to consider the oscillation (Definition 1.3.5).

***2.5.17. Proposition.***

*Let $f : [a, b] \to \mathbb{R}$ be a bounded function. The following statements are equivalent:*

*(i) $f$ is Riemann integrable;*

*(ii) the set of discontinuity points of $f$ is a nullset for Lebesgue measure.*

*Proof.* Let $N = \{x \in [a, b] : \ f \ \text{ is discontinuous at } x\}$. If we let $N_c = \{x : o(f, x) \geq c\}$ then $N = \bigcup_n N_{1/n}$. Each $B_{1/n}$ is compact by Lemma 1.3.6.

(i) $\implies$ (ii) Fix $\varepsilon > 0$ and $n \in \mathbb{N}$. By Exercise 1.3.7 there exists a partition $P = \{x_1, \ldots, x_h\}$ of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon/n$. Let $x_{r_1}, \ldots, x_{r_s} \subset P$ be those points in $P$ such that $[x_{r_j-1}, x_{r_j}] \cap B_{1/n} \neq \varnothing$. Then $M_{r_j}(f, P) - m_{r_j}(f, P) \geq 1/n$ by definition of upper and lower integral. Therefore

$$\begin{aligned}\sum_{j=1}^{s} x_{r_j} - x_{r_j-1} &\leq n \sum_{j=1}^{s} \left[M_{r_j}(f, P) - m_{r_j}(f, P)\right] (x_{r_j} - x_{r_j-1}) \\ &\leq n \big(U(f, P) - L(f, P)\big) < \varepsilon.\end{aligned}$$

As $N_{1/n} \subset \bigcup_{j=1}^{s} [x_{r_j-1}, x_{r_j}]$, we have shown that $m(N_{1/n}) < \varepsilon$. And $\varepsilon$ was arbitrary, so $m(N_{1/n}) = 0$. Hence $N$ is a countable union of nullsets, and therefore a nullset itself.

(ii) $\implies$ (i) Fix $\varepsilon > 0$. Now the hypothesis is that $N$ is a nullset, and so $N_\varepsilon$ is also a nullset. By the definition of Lebesgue measure and compactness (Lemma 1.3.6) there exist finitely many intervals $[x_1, y_1], \ldots, [x_s, y_s]$ such that $N_\varepsilon \subset \bigcup_j (x_j, y_j)$ and $\sum_{j=1}^{s} (y_j - x_{j-1}) < \varepsilon$. Let $P = \{x_1, y_1, x_2, y_2, \ldots, x_s, y_s\}$. By hypothesis there exists $c > 0$ with $|f(x)| \leq c$ for all $x$. We have

$$\sum_{j=1}^{s} \big[M_{y_j}(f, P) - m_{y_j}(f, P)\big](y_j - x_j) \leq 2c \sum_{j=1}^{s} (y_j - x_j) = 2c\varepsilon.$$

For $x \in [y_j, x_{j+1}]$ we have that $x \notin N_\varepsilon$ so $o(f, x) < \varepsilon$. This means that there exists an interval $I_x = (a_x, b_x)$ with $x \in I_x$ and $M_{I_x}(f) - m_{I_x}(f) < \varepsilon$. By compactness, there exist finitely many $x_{j,1}, \ldots, x_{j,s_j}$ such that $[y_j, x_{j+1}] \subset \bigcup_{r=1}^{s_j} (a_{x_j}, b_{x_j})$. Let $Q = \{a_{x_{j,r}}, b_{x_{j,r}} : j = 1, \ldots, s; \ r = 1, \ldots, s_j\}$. By construction, $U(f, Q) - L(f, Q) < \varepsilon(b - a)$. It follows that

$$U(f, P \cup Q) - L(f, P \cup Q) < 2c\varepsilon + \varepsilon(b - a).$$

As this can be done for any $\varepsilon > 0$, it follows that $f$ is integrable by Exercise 1.3.7. $\square$

### 2.5.3. Exercises.

(2.5.1) Show, without using the convergence theorems, that if $\mu$ is the counting measure on $\mathbb{N}$, for a non-negative sequence $\{a_n\}$ we have

$$\int_{\mathbb{N}} a\, d\mu = \sum_{n=1}^{\infty} a_n.$$

(2.5.2) Show, using the convergence theorems, that if $\mu$ is the counting measure on $\mathbb{N}$, for a non-negative sequence $\{a_n\}$ we have

$$\int_{\mathbb{N}} a\, d\mu = \sum_{n=1}^{\infty} a_n.$$

(2.5.3) Show that if $f \geq 0$ a.e. and $\displaystyle\int_X f\, d\mu = 0$, then $f = 0$ a.e. Conclude that if $f$ is measurable and $\displaystyle\int_X |f|\, d\mu = 0$, then $f = 0$ a.e.

(2.5.4) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f : X \to \overline{\mathbb{R}}$. Let $\{E_n\} \subset \mathcal{A}$ be a pairwise disjoint sequence. Let $E = \bigcup_n E_n$. Show that, if the integral on the left exists,

$$\int_E f\, d\mu = \sum_n \int_{E_n} f\, d\mu.$$

(2.5.5) Let $X$ be any set, $\Sigma = \mathcal{P}(X)$, $x_0 \in X$. Consider the Dirac measure (cfr. Exercise 2.3.6) $\delta : \Sigma \to [0, \infty]$ by

$$\delta(A) = \begin{cases} 1, & x_0 \in A \\ 0, & x_0 \notin A \end{cases}$$

Show that for any $f : X \to R$,

$$\int_X f\, d\delta = f(x_0).$$

(2.5.6) Let $f : [0, \infty) \to \mathbb{C}$ be integrable with respect to Lebesgue measure, and such that

$$\int_{[0,x]} f\, dm = 0, \qquad x > 0.$$

Show that $f = 0$ a.e. *(This can be certainly done with the techniques from this chapter, but the argument might not be all that direct)*

(2.5.7) *(Change of variable)* Let $(X, \mathcal{A}, \mu)$ be a measure space and $Y$ a topological space. Let $g : X \to Y$ measurable. Let $\nu(E) = \mu(g^{-1}(E))$. Show that $\nu$ is a measure on $\mathcal{B}(Y)$ and that

$$\int_Y f\, d\nu = \int_X f \circ g\, d\mu \tag{2.23}$$

for all $f : Y \to \overline{\mathbb{R}}$ which are Borel measurable and such that either integral exists.

(2.5.8) Consider the interval $[0,1]$ with Lebesgue measure. Let $g : [0,1] \to \mathbb{C}$ measurable and such that $\int_{[0,1]} |g| < \infty$. Prove that the function

$$\gamma : s \longmapsto \int_0^s |g|$$

is continuous.

(2.5.9) *(Change of variable, II)* Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to [0,\infty]$ measurable. Define $\nu : \mathcal{A} \to [0,\infty]$ by

$$\nu(E) = \int_E f \, d\mu.$$

(i) Show that $\nu$ is a measure on $(X, \mathcal{A})$ and that $\nu(E) = 0$ whenever $E \in \mathcal{A}$ with $\mu(E) = 0$.

(ii) Show that for any measurable function $g : X \to \overline{\mathbb{R}}$,

$$\int_X g \, d\nu = \int_X gf \, d\mu \tag{2.24}$$

if the left integral exists.

(2.5.10) *(Change of variable, III)* Use Exercises 2.5.7, 2.5.9 and 1.8.19 to show that if $g : [a,b] \to [c,d]$ is continuously differentiable then (2.23) gives the usual calculus change of variable (substitution) formula

$$\int_{g(a)}^{g(b)} f(t) \, dt = \int_a^b f(g(t)) \, g'(t) \, dt.$$

(2.5.11) Use Dominated Convergence and Exercise 1.3.8 to show that if $f : \mathbb{R} \to [0,\infty)$ is piecewise continuous with jump discontinuities, then with respect to Lebesgue measure

$$\int_{[a,b]} f \, dm = \int_a^b f(t) \, dt.$$

(2.5.12) Show that a linear combination of integrable functions is integrable.

(2.5.13) Compute $\displaystyle\int_{[0,1]} \alpha \, dm$, where $\alpha$ is Cantor's ternary function.

(2.5.14) Consider the function $f : [0,1] \to \mathbb{R}$ given by the series

$$f(x) = \sum_{n=0}^{\infty} \frac{x}{(1+x)^n}.$$

(i) Show that $f(x) = x + 1$ if $x \in (0,1]$, and that $f(0) = 0$.

(ii) Does the series converge uniformly on $[0,1]$?

(iii) Is is true that

$$\int_0^1 \left( \sum_{n=0}^{\infty} \frac{x}{(1+x)^n} \right) dx = \sum_{n=0}^{\infty} \int_0^1 \frac{x}{(1+x)^n} \, dx\,?$$

(2.5.15) Prove that

$$s = \lim_{k\to\infty} \sum_{n=1}^{\infty} \exp\left( -n + \frac{k}{n}\, e^{-k/n} \right)$$

exists and find its value.

(2.5.16) Find

$$\lim_{n\to\infty} \int_0^n x^{1/n}\, e^{-x}\, dx.$$

(2.5.17) Fix $\alpha, \beta > 0$.

(i) Show that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\alpha + n\beta} = \int_0^1 \frac{x^{\alpha-1}}{1+x^{\beta}}\, dx.$$

(ii) Show that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log 2.$$

(iii) Show that

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}.$$

(2.5.18) If $a_{n,m} \geq 0$ for all $n, m \in \mathbb{N}$, use Monotone Convergence to show that

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m}.$$

Show by example that the equality can fail in general.

(2.5.19) Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{X_n\} \subset \mathcal{A}$ a non-decreasing sequence such that $X = \bigcup_n X_n$. Show that, for any $f \geq 0$ measurable,

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_{X_n} f\, d\mu.$$

(2.5.20) Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{C}$ integrable and such that

$$\left| \int_X f\, d\mu \right| = \int_X |f|\, d\mu. \tag{2.25}$$

Show that there exists $\alpha \in \mathbb{C}$, with $|\alpha| = 1$ and $g : X \to [0, \infty)$ integrable such that $f = \alpha g$ a.e.

(2.5.21) Let $(X, \mathcal{A}, \mu)$ be a measure space. If $f : X \to \mathbb{C}$ is measurable and $E \in \mathcal{A}$ with $0 < \mu(E) < \infty$, show that

$$\frac{1}{\mu(E)} \int_E f \, d\mu \in \overline{\mathrm{conv}} f(E),$$

the closed convex hull of $f(E)$; that is, the average of $f$ is a limit of convex combinations of values of $f$.

(2.5.22) Let $(X, \mathcal{A}, \mu)$ be a measure space such that $\mu(X) < \infty$. Let $f : X \to \mathbb{C}$ be integrable, and $S \subset \mathbb{C}$ closed, and such that

$$\frac{1}{\mu(E)} \int_E f \, d\mu \in S, \qquad \text{whenever } E \in \mathcal{A},\ \mu(E) > 0.$$

Show that $f(x) \in S$ a.e.

(2.5.23) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $\{f_n\}$ a sequence of measurable functions such that $f_n \to 0$ uniformly. Does this imply that $\displaystyle\int_X f_n \, d\mu \to 0$?

(2.5.24) Let $f : X \to [0, \infty]$ measurable, with $c = \displaystyle\int_X f \, d\mu < \infty$, and $\alpha > 0$. Show that

$$\lim_{n\to\infty} \int_X n \log\left[1 + \left(\frac{f}{n}\right)^\alpha\right] d\mu = \begin{cases} \infty, & 0 < \alpha < 1 \\ c, & \alpha = 1 \\ 0, & 1 < \alpha < \infty \end{cases}$$

(2.5.25) Let $(X, \mathcal{A}, \mu)$ be a measure space, and $f : X \to [0, 1]$ integrable. Find

$$\lim_{n\to\infty} \int_X f^n \, d\mu.$$

(2.5.26) Let $(X, \mathcal{A})$ be a measurable space and $f : X \to [0, \infty]$ measurable. Show that

$$\int_0^\infty 1_{f^{-1}[t,\infty)}(s) \, dt = f(s), \qquad s \in X. \tag{2.26}$$

This is sometimes called the *layer-cake representation* of $f$.

(2.5.27) Find a sequence $\{f_n\}$ of continuous functions $f_n : [0, 1] \to \mathbb{R}$ such that

- $0 \le f_n \le 1$;
- for each $x \in [0, 1]$, $\lim_n f_n(x)$ does not exist;
- $\displaystyle\lim_n \int_0^1 f_n = 0$.

(2.5.28) Find a sequence of continuous functions $f_n : [0, 1] \to [0, \infty)$ such that

- $\displaystyle\lim_n f_n(x) = 0$ for all $x$;
- $\displaystyle\lim_n \int_0^1 f_n = 0$;

- $\lim_n \int_0^1 \sup f_n = \infty$.

*(note that these functions fail the hypotheses of DCT but they do satisfy the conclusion)*

## 2.6. Some More Topology

We will now develop several results about continuous functions on locally compact spaces. Uryson's Lemma and Tietze's Extension Theorem below will prove invaluable.

**2.6.1. Definition.**

*Given a locally compact Hausdorff space $T$, the **support** of $f : T \to \mathbb{C}$ is the set*

$$\operatorname{supp} f = \overline{\{t \in T : \ f(t) \neq 0\}}.$$

*The function $f$ is said to **vanish at infinity** if given $\varepsilon > 0$ there exists $K \subset T$, compact, with $|f(t)| < \varepsilon$ for all $t \in T \setminus K$.*

*We denote by $C(T)$ the space of continuous functions $f : T \to \mathbb{C}$, by $C_0(T)$ the space of continuous functions $f : T \to \mathbb{C}$ vanishing at infinity, and by $C_c(T)$ the space of continuous function $f : T \to \mathbb{C}$ with compact support.*

*On these spaces we consider the **supremum norm***

$$\|f\|_\infty = \sup\{|f(t)| : \ t \in T\}.$$

In the particular case where $T$ is compact, $C_0(T)$ and $C_c(T)$ coincide with $C(T)$. When the context is clear, it is common to drop the subindex for the norm $\|\cdot\|_\infty$.

The norm induces a metric via $d(f, g) = \|f - g\|_\infty$, so both $C_c(T)$ and $C_0(T)$ become metric spaces.

***2.6.2. Proposition.***

*Let $T$ be a topological space. Then $C_c(T)$ is dense in $C_0(T)$.*

*Proof.* The inclusion $C_c(T) \subset C_0(T)$ is automatic from the definition. Fix $g \in C_0(T)$ and $\varepsilon > 0$. Let $\gamma : \mathbb{C} \to \mathbb{C}$ be given by

$$\gamma(z) = \begin{cases} 0, & |z| \leq \varepsilon \\ 2(|z| - \varepsilon)\,\frac{z}{|z|}, & \varepsilon \leq |z| \leq 2\varepsilon \\ z, & |z| > 2\varepsilon \end{cases}$$

This function is continuous. Let $g_0 = \gamma \circ g$. Then $g_0$ is continuous and $\operatorname{supp} g_0 = \{t : \ |g(t)| \geq \varepsilon\}$ is compact by Exercise 2.6.2. When $|g(t)| \geq 2\varepsilon$ we have $g_0(t) = g(t)$; when $|g(t)| \leq \varepsilon$ we have $g_0(t) = 0$; and when $\varepsilon \leq |g(t)| \leq 2\varepsilon$, we have $|g_0(t)| \leq 2\varepsilon$. It follows that $\|g - g_0\|_\infty \leq 4\varepsilon$. As $\varepsilon$ was arbitrary, this shows that $C_c(T)$ is dense in $C_0(T)$. □

Towards proving the big theorems, we first show that compactness improves Hausdorff separation.

**2.6.3. Lemma.**

*Let $T$ be a Hausdorff space, $K \subset T$ compact, and $t \in T \setminus K$. Then there exist disjoint open sets $V, W \subset T$ with $K \subset V$ and $t \in W$.*

*Proof.* For each $x \in K$, by the Hausdorff property there exist open sets $V_x$ and $W_x$, disjoint, with $x \in V_x$ and $t \in W_x$. We have $K \subset \bigcup_x V_x$, so by compactness there exist $x_1, \ldots, x_m$ with $K \subset \bigcup_{j=1}^m V_{x_j}$. Let $V = \bigcup_{j=1}^m V_{x_j}$, $W = \bigcap_{j=1}^m W_{x_j}$. By construction, both are open and $V \cap W = \varnothing$, $K \subset V$, and $t \in W$. □

**2.6.4. Lemma.**

*Let $T$ be locally compact Hausdorff, and $K \subset V \subset T$ with $K$ compact and $V$ open. Then there exists $V' \subset T$, open and with compact closure, such that*

$$K \subset V' \subset \overline{V'} \subset V.$$

*Proof.* For each $t \in K$ there exists, since $T$ is locally compact, $V_t$ open with $t \in V_t$ and $\overline{V_t}$ compact. Then $K \subset \bigcup_t V_t$ and, by compactness, there exist $t_1, \ldots, t_r$ with $K \subset L = \bigcup_{j=1}^r V_{t_j}$. Note that $L$ is open with compact closure. If $V = T$, we take $V' = L$.

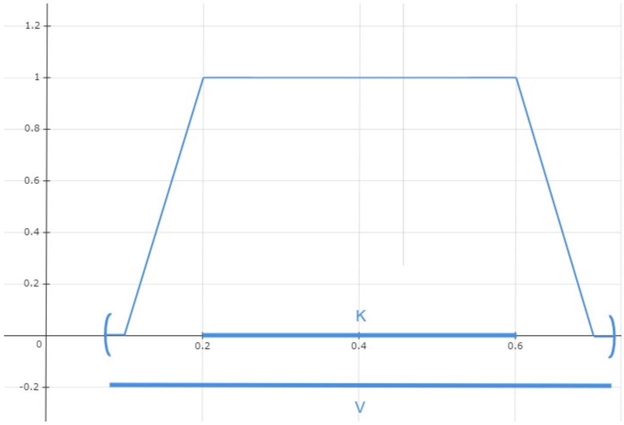
Otherwise, if $V \subsetneq T$, by Lemma 2.6.3 for each $t \in T \setminus V$ there exists an open set $W_t$ with $K \subset W_t$ and $t \notin \overline{W_t}$. Then the collection $\{(T \setminus V) \cap \overline{L} \cap \overline{W_t}\}_{t \in T \setminus V}$ consists of compact sets and has empty intersection (since $t \notin \overline{W}_t$).

By Proposition 1.8.19 there exist $t_1, \dots, t_s$ such that

$$(T \setminus V) \cap \overline{L} \cap \overline{W_{t_1}} \cap \cdots \cap \overline{W_{t_s}} = \varnothing. \tag{2.27}$$

Take $V' = L \cap W_{t_1} \cap \cdots \cap W_{t_s}$. This is open, $K \subset V'$, and by (2.27) we have $\overline{V}' \subset \overline{L} \cap \overline{W_{t_1}} \cap \cdots \cap \overline{W_{t_s}} \subset V$; the first inclusion shows that $\overline{V}'$ is compact. $\square$

One should admire the wonder of Urysohn's Lemma below. When $T = \mathbb{R}$, $K \subset \mathbb{R}$ is a closed interval, and $V \supset K$ is an open interval, it is simply saying that we can get a continuous function that is 1 on the compact set $K$, and it becomes zero outside of $V$:




Now, how would you do this when $K$ and $V$ are not intervals? Not worth it, but probably doable. What about for an arbitrary compact subset $K \subset V$ with $V \subset \mathbb{R}^{37}$ open? And, what about on an abstract topological space? Magic, that's the answer; or said in another way, Urysohn's Lemma.

**2.6.5. Theorem. (Urysohn's Lemma)**

*Let $T$ be a locally compact Hausdorff space. If $K \subset V \subset T$, with $V$ open and $K$ compact, there exists $f : T \to [0,1]$, continuous with compact support, with $f|_K = 1$ and $\operatorname{supp} f \subset V$.*

*Proof.* As we have seen before (Exercise 2.2.4, to be precise), the dyadic numbers $k/2^n$, with $n \in \mathbb{N}$ and $k = 1, \dots, 2^n - 1$, are dense in $[0,1]$. By applying Lemma 2.6.4 twice, there exist $V_{0,0}, V_{1,0} \subset T$ open, with compact closure, and such that

$$K \subset V_{1,0} \subset \overline{V_{1,0}} \subset V_{0,0} \subset \overline{V_{0,0}} \subset V.$$

Next we proceed inductively. Suppose we have open sets $V_{0,n}, \dots, V_{2^n,n}$ with

$$V_{1,0} = V_{2^n,n} \subset \overline{V_{2^n,n}} \subset V_{2^n-1,n} \subset \overline{V_{2^n-1,n}} \subset \cdots$$

$$\subset V_{1,n} \subset \overline{V_{1,n}} \subset V_{0,n} = V_{0,0}$$

and $V_{2k,n} = V_{k,n-1}$. Let $V_{2k,n+1} = V_{k,n}$, $k = 0, \ldots, 2^n$, and use Lemma 2.6.4 to obtain $V_{2k+1,n+1}$ open with compact closure such that

$$\overline{V_{2k+2,n+1}} \subset V_{2k+1,n+1} \subset \overline{V_{2k+1,n+1}} \subset V_{2k,n+1}, \qquad k = 0, \ldots, 2^n - 1.$$

The inclusion order of the sets $V_{k,n}$ corresponds with the order of the dyadic numbers $\frac{k}{2^n}$. Indeed, we associate $V_{k,n}$ with $\frac{k}{2^n}$. We have $V_{2k,n+1} = V_{k,n}$ coherent with $\frac{2k}{2^{n+1}} = \frac{k}{2^n}$. So in the initial step we created $V_{1,0}, V_{0,0}$, corresponding to 1 and 0. In the next step, when $n = 1$, we repeated $V_{2,1} = V_{1,0}$ and $V_{0,1} = V_{0,0}$, and create the new set $V_{1,1}$, which corresponds to $\frac{2}{2^2} = \frac{1}{2}$. So the second step mimics $1, \frac{1}{2}, 0$. For $n = 2$ we will mimic $1, \frac{3}{4}, \frac{1}{2}, \frac{1}{4}, 0$; and so on.

Now define $f(t) = 0$ for $t \notin V_{0,0}$, and

$$f(t) = \sup\left\{\frac{k}{2^n} : \ t \in V_{k,n}\right\}, \quad t \in V_{0,0}.$$

Fix $t \in V_{0,0}$ and choose $\varepsilon > 0$; fix $n \in \mathbb{N}$ such that $2^n > \frac{1}{\varepsilon}$. As the dyadic numbers $\frac{k}{2^n}$ for fixed $n$ partition the interval $[0, 1]$, there exists $k$ such that $\frac{k}{2^n} \leq f(t) < \frac{k+1}{2^n}$; so $|f(t) - \frac{k}{2^n}| < \frac{1}{2^n}$ and $t \in V_{k,n}$, $t \notin \overline{V_{k+1,n}}$ (if $t \in \overline{V_{k+1,n}}$ then $t \in V_{2^m(k+1)-1,n+m}$ for all $m$, which implies $f(t) \geq \frac{k+1}{2^n}$). Let $W = V_{k,n} \cap (T \setminus \overline{V_{k+1,n}})$, which is an open neighbourhood of $t$. If $s \in W$, then $\frac{k}{2^n} \leq f(s) < \frac{k+1}{2^n}$, and

$$|f(t) - f(s)| \leq \frac{1}{2^n} < \varepsilon.$$

Thus $f$ is continuous on $V_{0,0}$. If $t \in T \setminus \overline{V_{0,0}}$ then $f(t) = 0$ and there exists $W \subset T \setminus \overline{V_{0,0}}$ open with $t \in W$, so $f = 0$ on $W$ and thus $|f(t) - f(s)| = 0 < \varepsilon$ for all $s \in W$.

Finally, when $t \in \overline{V_{0,0}} \setminus V_{0,0}$, we have $f(t) = 0$. Fix $k, n$ with $\frac{k}{2^n} < \varepsilon$. Define $W = T \setminus \overline{V_{k,n}}$. Then $W$ is open, $t \in W$, and for any $s \in W$ we have that $f(s) \leq \frac{k}{2^n} < \varepsilon$. Then $|f(t) - f(s)| = f(s) < \varepsilon$.

So $f$ is continuous, $f|_K = 1$ (recall that $K \subset V_{1,0}$), $\operatorname{supp} f \subset \overline{V_{0,0}}$ which is compact, and $\operatorname{supp} f \subset V$. □

A careful examination of the proof of Urysohn's Lemma shows that the crucial tool is Lemma 2.6.4. This means that in any situation where $V'$ can be produced as in Lemma 2.6.4, the proof of Urysohn's Lemma will work. In particular, if we remove the word "compact" from the statement of Urysohn's Lemma, what one needs is a topological space where given $K \subset V$ with $K$ closed and $V$ open, there exists $V'$ open with $K \subset V' \subset \overline{V'} \subset V$; this is equivalent to a *normal* topological space (Exercise 1.8.6). Recall that a topological space $X$ is *normal* if given any two disjoint closed sets $C_1, C_2$, there exist disjoint open sets $V_1 \supset C_1$ and $V_2 \supset C_2$. This gives us

**2.6.6. Theorem. (General Urysohn's Theorem)**

*Let $X$ be a topological space. The following statements are equivalent:*

*(i) $X$ is normal;*

*(ii) given $K \subset V \subset X$ with $K$ closed and $V$ open, there exists $f : X \to [0,1]$, continuous, with $f|_K = 1$ and $f|_{X\setminus V} = 0$.*

*Proof.* (i) $\implies$ (ii) The closed sets $K$ and $X \setminus V$ are disjoint. By normality, there exist $V', V'' \subset X$, open and disjoint, with $K \subset V'$ and $X \setminus V \subset V''$. Because $V' \subset X \setminus V''$ and the latter is closed, $\overline{V'} \subset X \setminus V''$. Then $K \subset V' \subset \overline{V'} \subset V$.

Now we can rehash the proof of Urysohn's Lemma with the above playing the role of Lemma 2.6.4.

(ii) $\implies$ (i) Let $C_1, C_2 \subset X$ be closed and disjoint. Applying the hypothesis to the inclusion $C_1 \subset X \setminus C_2$, we get a continuous $f : X \to [0,1]$ such that $f|_{C_1} = 1$, $f|_{C_2} = 0$. Let $V_1 = f^{-1}(1/2, 2)$ and $V_2 = f^{-1}(-1, 1/2)$. Then $V_1, V_2$ are open, disjoint, with $C_1 \subset V_1$ and $C_2 \subset V_2$. $\square$

Though this is really not a concern in our context, not every locally compact Hausdorff space is normal. The locally compact version in Theorem 2.6.5 only produces the $f$ for $K$ compact, and not for arbitrary closed as in the normal case.

As mentioned in Proposition 1.8.10, one formulation of continuity for a function $f$ is that for every convergent net $\{t_j\}$ with $t_j \to t$, we have $f(t_j) \to f(t)$. For good enough spaces, the converse holds; that is, we can test convergence of nets by evaluating on (all) continuous functions. Note that metric spaces are normal.

**2.6.7. Corollary.**

*Let $T$ be a locally compact Hausdorff or normal space, and let $\{t_j\} \subset T$. Then $t_j \to t$ if and only if $f(t_j) \to f(t)$ for every $f \in C(T)$.*

*Proof.* If $t_j \to t$ and $f \in C(T)$, then $f(t_j) \to f(t)$ by definition of continuity .

Conversely, suppose that $t_j \not\to t$. Then there exists an open neighbourhood $V$ of $t$ and a subnet $\{t_{j_k}\}$ such that $t_{j_k} \not\in V$ for all $k$. By Urysohn's Lemma or Theorem 2.6.6 there exists $f \in C(T)$ such that $f(t) = 1$ and $f|_{V^c} = 0$. Then $f(t_{j_k}) = 0$ for all $k$, so $|f(t) - f(t_{j_k})| = 1$ for all $k$ and thus $f(t_{j_k}) \not\to f(t)$. $\square$

When $T$ is locally compact Hausdorff in Corollary 2.6.7, we can restrict the family of functions we use to those in $C_c(T)$; this is due to the fact that Urysohn's Lemma produces $f \in C_c(T)$.

*2.6.8. Remark*. The hypothesis in Corollary 2.6.7 is necessary. The problem in an arbitrary topological space is that there might be not enough continuous functions to test the convergence of the net; that is precisely where local compactness or normality, via Urysohn, comes to the rescue. For an extreme example of what can fail, consider the set $X = \{1, 2, 3\}$ with the topology $\{\varnothing, X, \{1, 2\}, \{2\}, \{2, 3\}\}$. This is of course not Hausdorff, and in fact $f : X \to \mathbb{R}$ is continuous if and only if $f$ is constant (Exercise 1.8.21). So we can consider the sequence $\{t_n\}$ with $t_n = 1 + (n \pmod 3)$. This sequence is not convergent, because for each of $1, 2, 3$ there is a neighbourhood such that the sequence goes out of the neighbourhood infinitely often. But as the only continuous functions are the constants, $f(t_n) \to f(t)$ for every continuous function $f : X \to \mathbb{R}$.

For an example where Urysohn fails in a Hausdorff space, by Theorem 2.6.6 any Hausdorff non-normal space will do. For instance consider $X = \mathbb{R}$ with the **K-topology**, which is the topology generated by the open intervals $(a, b)$, such that $a < b$, together with the sets $(a, b) \setminus K$, where

$$K = \left\{\frac{1}{n} : \ n \in \mathbb{N}\right\}.$$

The feature of this topology is that $K$ is closed. Indeed, this amounts to showing that $\mathbb{R} \setminus K$ is open; this is the trivial fact that $\mathbb{R} \setminus K = \bigcup_n \big[(-n, n) \setminus K\big]$, which is a union of open sets by definition. The topology is Hausdorff because points can be separated as in the usual topology in $\mathbb{R}$. The counterexample comes from considering the sets $\{0\}$ and $K$. Both are closed, and they are disjoint. If $\{0\} \subset V$ with $V$ open and $V \cap K = \varnothing$, there has to exist $a > 0$ such that $(-a, a) \setminus K \subset V$. Fix $n > 1/a$. If $W$ is open and $K \subset W$, there has to exist $r > 0$ such that $\left(\frac{1}{n} - r, \frac{1}{n} + r\right) \subset W$. As $a > \frac{1}{n} - r$, $V \cap W \neq \varnothing$; that is, 0 and $K$ cannot be separated by open sets. In particular, there cannot exist $f : \mathbb{R} \to \mathbb{R}$ continuous with $f(0) = 1$ and $f|_K = 0$, as that would allows us to produce disjoint $V, W$ with $\{0\} \subset V$ and $K \subset W$. In a similar vein, if $f$ is continuous for the $K$-topology then it is continuous for the usual topology on $\mathbb{R}$, as the $K$-topology is finer. So if we consider the sequence $\{\frac{1}{n}\}_{n\in\mathbb{N}}$, then $f(\frac{1}{n}) \to f(0)$ for all $K$-continuous $f$, even though $\frac{1}{n} \not\to 0$ in the $K$-topology. This provides a counterexample promised on page 62.

For Tietze's Extension Theorem below, we follow the proof from [Rud87, Theorem 20.4].

**2.6.9. Theorem.** *(Tietze's Extension Theorem)*

*Let $T$ be locally compact Hausdorff (or normal) space, and let $K \subset T$ be compact (in the locally compact case) or closed (in the normal case). For each complex-valued $f \in C(K)$, there exists $\tilde{f} \in C(T)$, of compact support in the locally compact case, and such that $\tilde{f}|_K = f$ and $\|\tilde{f}\|_\infty = \|f\|_\infty$.*

*Proof.* Assume first that $f$ is real, with $-1 \le f \le 1$. By Lemma 2.6.4 (or by normality, see the comment before Theorem 2.6.6) there exists $W$ open with compact closure and $K \subset W$.

Let

$$K^+ = \{t \in K : \ f(t) \ge \frac{1}{3}\}, \quad K^- = \{t \in K : \ f(t) \le -\frac{1}{3}\}.$$

Then both sets are compact, $K^+ \cup K^- \subset W$, and $K^+ \cap K^- = \varnothing$. Using Urysohn's Lemma, construct compactly supported $g_+, g_- \in C_c(T)$ such that $\operatorname{supp} g_\pm \subset W$, $0 \le g_+, g_- \le 1$, $g_+|_{K^+} = 1$, $g_+|_{K^-} = 0$, $g_-|_{K^+} = 0$, $g_-|_{K^-} = 1$. Let $f_1 = \frac{1}{3}(g_+ - g_-)$. Then $f_1 \in C_c(T)$ and $-\frac{1}{3} \le f_1 \le \frac{1}{3}$; moreover, $f_1|_{K^+} = \frac{1}{3}$, $f_1|_{K^-} = -\frac{1}{3}$. On $K^+$ we have that $|f - f_1| = |f - \frac{1}{3}| \le \frac{2}{3}$; similarly, on $K^-$ we have $|f - f_1| \le \frac{2}{3}$. And on $K \setminus (K^+ \cup K^-)$ we have $|f| \le \frac{1}{3}$ and $|f_1| \le \frac{1}{3}$. Thus

$$|f - f_1| \le \frac{2}{3}, \quad |f_1| \le \frac{1}{3}, \qquad \text{on } K. \tag{2.28}$$

Now we repeat the process inductively. Assume that we have $f_1, \ldots, f_n \in C_c(T)$, with $\operatorname{supp} f_j \subset W$ for all $j$, and such that

$$|f - f_1 - \cdots - f_n| \le \left(\frac{2}{3}\right)^n \text{ on } K, \quad |f_n| \le \frac{1}{3}\left(\frac{2}{3}\right)^{n-1} \text{ on } T. \tag{2.29}$$

Apply the construction above to $(\frac{3}{2})^n(f - f_1 - \cdots - f_n)$ to obtain $f'_{n+1} \in C_c(T)$, with $\subset f_{n+1} \subset W$, and such that

$$\left|\left(\frac{3}{2}\right)^n (f - f_1 - \cdots - f_n) - f'_{n+1}\right| \le \frac{2}{3} \text{ on } K, \quad |f'_{n+1}| \le \frac{1}{3} \text{ on } T.$$

So if we let $f_{n+1} = (\frac{2}{3})^n f'_{n+1}$, we obtain

$$|f - f_1 - \cdots - f_{n+1}| \le \left(\frac{2}{3}\right)^{n+1} \text{ on } K, \quad |f_{n+1}| \le \frac{1}{3}\left(\frac{2}{3}\right)^n \text{ on } T. \tag{2.30}$$

Now let $\tilde{f} = \sum_{j=1}^\infty f_j$. The inequalities in (2.30) show that the series converges to $f$ on $K$, and that it converges uniformly on $T$ (Exercise 2.6.5); so $\tilde{f}$ is continuous, it coincides with $f$ on $K$, and its support is contained in $\overline{W}$.

For arbitrary real $f$, we can do the above for $f/\|f\|_\infty$ (note that $f \in C(K)$, so $f$ is bounded). In general, write $f = f_1 + if_2$ with $f_1, f_2$ real, and extend each to get an extension $\tilde{f}$ that coincides with $f$ on $K$, possibly with

$\|\tilde{f}\|_\infty > \|f\|_\infty$. To tame the norm we consider the function $\gamma : \mathbb{C} \to \mathbb{C}$ given by

$$\gamma(z) = \begin{cases} z, & |z| \leq \|f\|_\infty \\ \frac{\|f\|_\infty z}{|z|}, & |z| > \|f\|_\infty \end{cases}$$

Then $\gamma$ is continuous, and if we replace the extension $\tilde{f}$ obtained above with $\gamma \circ \tilde{f}$, all needed properties are preserved, and $\|\gamma \circ \tilde{f}\|_\infty = \|f\|_\infty$. $\square$

*2.6.10. Remark*. Urysohn's Lemma is a particular case of Tietze's Extension Theorem. Indeed, if $T$ is a locally compact Hausdorff or normal space and $K \subset V$ with $K$ compact and $V$ open, the property allows us to find $W$ open with $\overline{W}$ compact if $T$ is locally compact, and $K \subset W \subset \overline{W} \subset V$. Applying Tietze on the topological space $\overline{W}$ to the function $f = 1$ on $K$ and $f = 0$ on $\overline{W} \setminus W$, as the set $K \cup (\overline{W} \setminus W)$ is closed (compact if $T$ is locally compact) by Tietze we get a function $\tilde{f}$ that extends $f$ and has closed (compact if $T$ is locally compact) support. If we further extend $\tilde{f}$ as 0 on $T \setminus \overline{W}$ we obtain a continuous function that is 1 on $K$ and with $\operatorname{supp} \tilde{f} \subset V$.

It turns out that while measurable functions can be crazy discontinuous, they are still fairly regular. That is the theme of Lusin's Theorem below. Pay attention to the fact that we are not *approximating* our measurable function, but rather we are showing that it agrees with a continuous function over a big chunk of the measure space. We prove two versions of Lusin's Theorem; first a general version where we show that in a good enough measure space it is always possible to remove a tiny part of the space and our measurable function is continuous on what remains. And then another version slightly more restrictive in its hypotheses but that in exchange produces a globally continuous function that restricts to our measurable function on a set of arbitrarily big measure.

Recall that in a measure space with finite measure, inner regularity by closed sets is equivalent to outer regularity (Exercise 2.3.20).

***2.6.11. Theorem. (Lusin, general version)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space with finite measure that also carries a Hausdorff topological structure, such that $\mathcal{B}(X) \subset \mathcal{A}$ and $\mu$ is outer regular. Let $Y$ be a second countable topological space. If $f : X \to Y$ is measurable then for every $\varepsilon > 0$ there exists $K \subset X$ closed, with $\mu(X \setminus K) < \varepsilon$, and $f|_K$ continuous.*

*Proof.* Let $\{V_n\}_{n\in\mathbb{N}}$ be a countable base for the topology of $Y$. Define, for each $n \in \mathbb{N}$,

$$E_n = f^{-1}(V_n), \qquad F_n = f^{-1}(Y \setminus V_n).$$

By the regularity of $\mu$ there exist closed sets $K_n \subset E_n$ and $H_n \subset F_n$ with

$$\mu(E_n \setminus K_n) < 2^{-n-1}\,\varepsilon, \qquad \mu(F_n \setminus H_n) < 2^{-n-1}\,\varepsilon$$

(this is where we use that $\mu(X) < \infty$, so that $E_n$ and $F_n$ have finite measure and outer regularity is equivalent to inner regularity by closed sets). Let

$$K = \bigcap_n (K_n \cup H_n).$$

Then $K$ is closed. We have, using that $E_n \cup F_n = X$ and $E_n \cap F_n = \varnothing$,

$$\begin{aligned}\mu(X \setminus (K_n \cup H_n)) &= \mu((E_n \cup F_n) \setminus (K_n \cup H_n))\\ &= \mu((E_n \setminus K_n) \cup (F_n \setminus H_n))\\ &\le 2^{-n-1}\,\varepsilon + 2^{-n-1}\,\varepsilon = 2^{-n}\,\varepsilon.\end{aligned}$$

Hence

$$\mu(X \setminus K) = \mu\Big(\bigcup_n (X \setminus (K_n \cup H_n))\Big) \le \sum_n \mu(X \setminus (K_n \cup H_n)) \le \sum_n 2^{-n}\,\varepsilon = \varepsilon.$$

To check that $f|_K$ is continuous it is enough to show that $f^{-1}(V_n) \cap K$ is relatively open in $K$ for each $n$. Equivalently, we want to show that $f^{-1}(Y \setminus V_n) \cap K$ is relatively closed in $K$. We have the obvious inclusion

$$H_n \cap K \subset f^{-1}(Y \setminus V_n) \cap K.$$

Suppose that $z \in f^{-1}(Y \setminus V_n) \cap K$. Since $K \subset K_n \cup H_n$ and these are disjoint, we have that $z \in K_n$ or $z \in H_n$. As $K_n \subset f^{-1}(V_n)$, it is impossible to have $z \in K_n$, and hence $z \in H_n$. Thus

$$f^{-1}(Y \setminus V_n) \cap K = H_n \cap K$$

is relatively closed in $K$ and $f|_K$ is continuous. □

The hypotheses in Theorem 2.6.11 (finite measure and outer regularity) cannot be relaxed (Exercises 2.6.3 and 2.6.4).

*2.6.12. Remark.* Lusin's Theorem 2.6.11 seems impressive—and it is—because very little is assumed. It is important to note, though, that the theorem is not saying that $f$ is continuous at every point of $K$, but rather that $f|_K$ is continuous. And these are (very) different things. An enlightening point of view is to consider $X = [0,1]$, and $f = 1_{\mathbb{Q}}$. This function is measurable and it is discontinuous everywhere. But restricted to $X \setminus \mathbb{Q}$, $f$ is constant and hence (very) continuous! The point of this next version of Lusin's Theorem is that with more restrictive hypotheses we can get $f$, up to a small set, as the restriction of an honest continuous function.

***2.6.13. Corollary. (Lusin, complex valued)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space with finite measure that also carries a locally compact Hausdorff (or normal) topological structure, such that $\mathcal{B}(X) \subset \mathcal{A}$ and $\mu$ is outer and inner regular. If $f : X \to \mathbb{C}$ is measurable then for every $\varepsilon > 0$ there exists $K \subset X$ compact, with $\mu(X \setminus K) < \varepsilon$, and $g \in C_c(X)$ such that $f|_K = g|_K$ and $\|f\|_\infty = \|g\|_\infty$.*

*Proof.* With inner regularity the proof of Theorem 2.6.11 now gives us $K$ compact (from the first choice of $K_n$ and $H_n$), and the local compactness allows us to use Tietze's Extension Theorem (2.6.9) to obtain $g \in C_c(X)$ with $g|_K = f|_K$ and $\|f\|_\infty = \|g\|_\infty$. $\square$

Even in the case of Corollary 2.6.13 the finiteness hypothesis cannot be relaxed (Exercise 2.6.4). It is worth mentioning that there is a very nice constructive proof of Lusin's theorem in [Rud87, Theorem 2.23]. We offer a sketch in Exercise 2.6.10.

***2.6.14. Corollary.***

*Under the same hypotheses as in Corollary 2.6.13, there exists a sequence $\{g_n\} \subset C_c(X)$ with $f = \lim_n g_n$ a.e.*

*Proof.* For each $n \in \mathbb{N}$ there exists, by Corollary 2.6.13, a function $g_n \in C_c(X)$ such that $\mu(\{f \neq g_n\}) < 2^{-n}$. Let $E_n = \{f \neq g_n\}$ and

$$B = \{x : \ |\{n : \ x \in E_n\}| = \infty\}, \qquad h = \sum_{n=1}^{\infty} 1_{E_n}.$$

Then $x \in B$ if and only if $h(x) = \infty$. We also have

$$\int_X h \, d\mu = \sum_{n=1}^{\infty} \mu(E_n) < \infty.$$

So $h < \infty$ a.e., which implies that $\mu(B) = 0$. Any $x \in X \setminus B$ can belong to only finitely many $E_n$; thus for $n$ big enough we have $f(x) = g_n(x)$. That is, on $X \setminus B$ we have $\lim_n g_n = f$. $\square$

Dini's Theorem below has nothing to do with measure, it is just a topological result; it says that if a monotone sequence of continuous functions has continuous limit on a compact set, then the convergence is uniform. Still this is the right place for it since it precludes Egorov's Theorem below.

**2.6.15. Theorem. (Dini)**

*Let $X$ be compact and $f_n, f : X \to \mathbb{R}$ continuous such that $f_n \nearrow f$ pointwise. Then the convergence is uniform.*

*Proof.* Fix $\varepsilon > 0$ and define

$$E_n = \{f - f_n < \varepsilon\} = (f - f_n)^{-1}(-\infty, \varepsilon).$$

As both $f_n$ and $f$ are continuous, each $E_n$ is open. The monotonicity of the sequence $\{f_n\}$ makes the sets $E_n$ also increasing, since $f - f_{n+1} \leq f - f_n$. The convergence of the sequence gives us that $X \subset \bigcup_n E_n$. By compactness, and since $\{E_n\}$ increases, there exists $m$ with $X \subset E_m$. By enlarging $m$ if needed, we may assume that $m > 1/\varepsilon$. Then for any $n \geq m$ we have

$$|f(x) - f_n(x)| = f(x) - f_n(x) = \big[f(x) - f_m(x)\big] + \big[f_m(x) - f_n(x)\big] < \varepsilon$$

for all $x \in X$ (since $f_m - f_n \leq 0$); that is, uniform convergence. □

For measurable functions, we can get uniform convergence up to small measure:

**2.6.16. Theorem. (Egorov)**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, and $E$ measurable with $\mu(E) < \infty$. Suppose that $\{f_n\}$ is a sequence of measurable functions with $f_n \to f$ pointwise a.e. on $E$. Then for every $\varepsilon > 0$ there exist $F \subset E$, measurable with $\mu(F) < \varepsilon$, and $f_n \to f$ uniformly on $E \setminus F$.*

*Proof.* By putting the nullset where the sequence does not converge to $f$ together with $F$, we may assume convergence everywhere. Define sets

$$E_n^m = \bigcap_{j=n}^{\infty} \Big\{x \in E : \ |f_j(x) - f(x)| < \frac{1}{m}\Big\}.$$

By construction, $E_1^m \subset E_2^m \subset \cdots$ for all $m$. As $f_n \to f$ on $E$, we have that $E \subset \bigcup_n E_n^m$ for all $m$. So this is an equality, and because the union is increasing we get by continuity of the measure that $\lim_n \mu(E \setminus E_n^m) = 0$ for each $m$. This allows us to choose, given $\varepsilon > 0$, an $n_0(m)$ such that

$$\mu(E \setminus E^m_{n_0(m)}) < \frac{\varepsilon}{2^m}.$$

Define $F = \bigcup_{m=1}^{\infty} (E \setminus E^m_{n_0(m)})$. The set $F$ is measurable, and

$$\mu(F) \leq \sum_{m=1}^{\infty} \mu(E \setminus E^m_{n_0(m)}) < \varepsilon.$$

If $x \in E \setminus F = \bigcap_{m=1}^{\infty} E^m_{n_0(m)}$, we can choose $m$ such that $m > 1/\varepsilon$, and then for any $n \geq n_0(m)$ we have $x \in E^m_n$ and so $|f_n(x) - f(x)| < \frac{1}{m} < \varepsilon$. That is, the convergence is uniform on $E \setminus F$. $\square$

We did not mention in the proof of Egorov's Theorem where we used that $\mu(E) < \infty$; identifying the right moment in the proof (Exercise 2.6.11) should help understanding. This is no minor matter, as Egorov's Theorem fails in general if $\mu(E) = \infty$. For instance consider $\mathbb{R}$ with Lebesgue measure, and $f_n = 1_{[n,n+1]}$. Then $f_n \to 0$, but $m(\{|0 - f_n| \geq 1\}) = \infty$ for all $n$.

**2.6.1. Exercises.**

(2.6.1) Show that If $X$ is a Hausdorff topological space and $K_1, K_2 \subset X$ are disjoint compact sets, there exist disjoint open sets $V_1, V_2$ with $K_1 \subset V_1$ and $K_2 \subset V_2$. Conclude that a compact Hausdorff space is normal.

(2.6.2) Let $T$ be a Hausdorff topological space and $f : T \to \mathbb{C}$ continuous. Show that the following statements are equivalent:

(i) $f$ vanishes at infinity;

(ii) for each $\varepsilon > 0$, the set $\{|f| \geq \varepsilon\}$ is compact.

(2.6.3) Show that if $X = [0, 1]$ with the usual topology, $\mathcal{A} = \mathcal{P}(X)$, and $\mu$ is the counting measure, there exists measurable $f : X \to \mathbb{C}$ such that the locally compact version of Lusin's Theorem fails. Discuss which hypothesis of the theorem is not satisfied in this example.

(2.6.4) Show that if $X = \mathbb{R}$ with the usual topology, $\mathcal{A} = \mathcal{B}(X)$ is the Borel $\sigma$-algebra, and $\mu$ is the Lebesgue measure, there exists measurable $f : X \to \mathbb{C}$ (taking nonzero values in a set of infinite measure) such that Lusin's Theorem fails.

(2.6.5) In the proof of Theorem 2.6.9, write the details of the argument that the inequalities in (2.30) show that the series converges to $f$ on $K$, and that it converges uniformly on $T$.

(2.6.6) Let $K \subset \mathbb{R}$ be closed. Is $K$ the support of a continuous real or complex-valued function? If it is, show how to construct such a function; if it isn't, describe which closed sets are supports of continuous functions. Does your answer apply to other topological spaces?

(2.6.7) Let $X$ be compact Hausdorff, and $f \in C(X)$. Suppose that $f(x) \neq 0$ for all $x \in X$. Show that $1/f \in C(X)$.

(2.6.8) Let $X$ be a locally compact Hausdorff space. Show that the following statements are equivalent:

(i) every $f \in C_0(X)$ is constant;

(ii) $X$ is a singleton.

(2.6.9) The goal of this exercise is to consider an alternative proof of Lusin's Corollary 2.6.13 in the case where $X$ is a metric space. Concretely, we want to show that if $f : X \to \mathbb{C}$ is measurable then for every $\varepsilon > 0$ there exists $K \subset X$ compact, with $\mu(X \setminus K) < \varepsilon$, and $g \in C(X)$ such that $g|_K = f|_K$ and $\|f\|_\infty = \|g\|_\infty$.

(i) Assume that $f$ is real-valued. Show that because $\mu$ is inner regular, the proof Theorem 2.6.11 can be repeated with "compact" in place of "closed".

(ii) Fix $\varepsilon > 0$ and let $K$ be the compact subset obtained in the modified version of Theorem 2.6.11; that is, $K \subset X$ is compact with $\mu(X \setminus K) < \varepsilon$ and such that $f|_K$ continuous. Conclude that $f$ is uniformly continuous.

(iii) Define the *modulus of continuity* $\omega$ of $f$, to be the function $\omega_1 : [0, \infty) \to [0, \infty)$ with

$$\omega_1(t) = \sup\{|f(y) - f(x)| : \ x, y \in K, \ d(x, y) \leq t\}.$$

Show that $|f(x) - f(y)| \leq \omega_1(d(x, y))$ for all $x, y \in K$, that $\omega_1(0) = 0$, and that $\omega$ is non-decreasing and continuous at 0.

(iv) Let

$$\omega(t) = \frac{1}{t} \int_t^{2t} \omega_1(s)\, ds.$$

Show $\omega$ is continuous, that we can define $\omega(0) = 0$ while maintaining continuity, and $|f(x) - f(y)| \leq \omega(d(x, y))$ for all $x, y \in K$.

(v) Define

$$g(x) = \inf\{f(y) + \omega(d(x, y)) : \ y \in K\}. \tag{2.31}$$

Show that $g$ is continuous and that $g|_K = f|_K$.

(vi) Show that if $f$ is complex-valued, a continuous function $g$ as above exists for $f$.

(vii) If $\|f\|_\infty < \infty$, show that $g$ can be replaced with $h \circ g$, in the sense that there exists a continuous function $h : \mathbb{C} \to \mathbb{C}$ such that $\|h \circ g\|_\infty = \|f\|_\infty$, and $h \circ g$ is still a continuous function that agrees with $f$ on $K$.

(2.6.10) This exercise sketches the constructive proof of Lusin's Theorem (Corollary 2.6.13) given in [Rud87, Theorem 2.24]. Properly, the statement to be proven is that in the situation of Corollary 2.6.13, when the measure space is complete and the measure $\mu$ is outer and inner regular, if there exists $A \in \mathcal{A}$ with $\mu(A) < \infty$ and $f|_{X\setminus A} = 0$, then for every $\varepsilon > 0$ there exists $g \in C_c(X)$ and $B \in \mathcal{A}$, such that $\mu(X \setminus B) < \varepsilon$, $f = g$ on $B$. If $\|f\|_\infty < \infty$ we can choose $g$ with $\|f\|_\infty = \|g\|_\infty$.

(i) Assume $A$ compact. Show that there exists $V$ open with $\overline{V}$ compact and $A \subset V$.

(ii) Show that if $0 \le f \le 1$ there exists a sequence $\{s_n\}$ of simple functions with $s_n \nearrow f$ uniformly, and such that

$$s_n(X) \subset \left\{\frac{m}{2^n},\ m \in \mathbb{N}\right\}$$

and

$$s_n(x) - s_{n-1}(x) \in \{0, 2^{-n}\},\ x \in X.$$

(iii) Define simple functions $\{t_n\}$ by

$$t_1 = s_1, \qquad t_n = s_n - s_{n-1}, \qquad n \in 1 + \mathbb{N}.$$

Show that, with $s_0 = 0$,

$$\sum_{n=1}^{\infty} t_n = f.$$

(iv) For each $n \in \mathbb{N}$, let $T_n = t_n^{-1}(\{2^{-n}\})$. Show that $T_n$ is measurable, and $T_n \subset A \subset V$ for all $n$.

(v) Show that for each $n$ there exist $K_n$ compact and $V_n$ open, with $K_n \subset T_n \subset V_n \subset V$ and with $\mu(V_n \setminus K_n) < \frac{\varepsilon}{2^n}$, and $g_n \in C_c(X)$ with $g_n|_{K_n} = 1$, $0 \le g \le 1$, and $\operatorname{supp} g_n \subset V_n$.

(vi) Define

$$g = \sum_{n=1}^{\infty} 2^{-n} g_n.$$

Show that $0 \le g_n \le 1$, $g \in C_c(X)$, and there exists $B \in \mathcal{A}$ with $\mu(X \setminus B) < \varepsilon$ and $f = g$ on $B$.

(vii) Extend the result to arbitrary $A$ measurable and arbitrary $f$.

(2.6.11) Determine where in the proof of Egorov's Theorem 2.6.16 is the condition $\mu(E) < \infty$ used.

(2.6.12) Let $(X, \mathcal{M}, \mu)$ be a positive measure space. We use the following notation, that will be introduced formally soon: $L^1(\mu)$ is the set of integrable functions; and $\|f\|_1 = \displaystyle\int_X |f|\, d\mu$. A set $\Phi \subset L^1(\mu)$ is said to be *uniformly integrable* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \int_E f \, d\mu \right| < \varepsilon \tag{2.32}$$

whenever $f \in \Phi$ and $\mu(E) < \delta$.

(i) Prove that every finite subset of $L^1(\mu)$ is uniformly integrable.

(ii) Prove the following convergence theorem of Vitali: if

  (a) $\mu(X) < \infty$;

  (b) $\{f_n\}$ is uniformly integrable;

  (c) $f_n(x) \to f(x)$ a.e.;

  (d) $|f(x)| < \infty$ a.e.;

  then $f \in L^1(\mu)$ and $\|f_n - f\|_1 \to 0$.
  *Suggestion:* Egorov.

(iii) Show that (ii) fails for the Lebesgue measure on $\mathbb{R}$, even if $\|f_n\|_1 \leq c$ for all $n$. So the finite-measure hypothesis cannot be omitted.

(iv) Show that for a measure space $X$ with finite measure, Vitali's Theorem (ii) implies the Dominated Convergence Theorem.

## 2.7. Product Measures

It is common to create a new object out of two others via the Cartesian product. This is done for many different structures, like sets, groups, rings, vector spaces, modules, etc. Besides the plain case of sets, in all other cases one needs to build the associated structure, and most of the time this is fairly straightforward. With measure spaces, we have constructed a measure structure on $\mathbb{R}$. On $\mathbb{R}^d$, we could adapt the construction of the Lebesgue outer measure, $\sigma$-algebra, and measure; this works, and it is instructive to think about the details for a while. We will not pursue said avenue because we will see two other ways of constructing Lebesgue measure on $\mathbb{R}^d$. The first one occurs in this section, where we will construct a measure space on the Cartesian product of two measure spaces; by iterating, this allows us to extend the construction to any finite Cartesian product of measure spaces.

**2.7.1. Definition.**

*Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$ be measurable spaces. A* ***measurable rectangle*** *is subset $A \times B \subset X \times Y$, with $A \in \mathcal{A}$, $B \in \mathcal{B}$. The* ***product $\sigma$-algebra*** *$\mathcal{A} \boxdot \mathcal{B}$ is the $\sigma$-algebra generated by the measurable rectangles. The product of $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ is the measurable space $(X \times Y, \mathcal{A} \boxdot \mathcal{B})$.*

*If $E \subset X \times Y$ and $x \in X$, $y \in Y$, the* ***$x$-section*** *and the* ***$y$-section*** *of $E$ are the sets $E_x \subset Y$ and $E^y \subset X$ given by*

$$E_x = \{y \in Y : \ (x, y) \in E\}, \qquad E^y = \{x \in X : \ (x, y) \in E\}.$$

There is no standard notation for the product $\sigma$-algebra $\mathcal{A} \boxdot \mathcal{B}$ (among others, $\times$ and $\otimes$ are used in the literature); we have chosen a symbol that is not usually used elsewhere.

For a rectangle $E = A \times B$ we have $E_a = B$, $E_b = A$ for all $a \in A$, $b \in B$; and $E_a = \varnothing$ when $a \notin A$, $E^b = \varnothing$ when $b \notin B$. The sections satisfy the natural relations

$$\Big(\bigcup_k E_k\Big)_x = \bigcup_k (E_k)_x, \qquad \Big(\bigcap_k E_k\Big)_x = \bigcap_k (E_k)_x, \qquad (E^c)_x = (E_x)^c$$

(Exercise 2.7.1).

***2.7.2. Proposition.***

*Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, and $E \in \mathcal{A} \boxdot \mathcal{B}$. Then, for every $x \in X$ and every $y \in Y$, the $x$-section and the $y$-section are measurable; that is, $E_x \in \mathcal{B}$, $E^y \in \mathcal{A}$.*

*Proof.* Fix $x \in X$. Consider the function $\pi_x : Y \to X \times Y$ given by $\pi_x(y) = (x, y)$. Let

$$\mathcal{R} = \{G \in \mathcal{A} \boxdot \mathcal{B} : \ \pi_x^{-1}(G) \in \mathcal{B}\}.$$

By Exercise 2.4.1, $\mathcal{R}$ is a $\sigma$-algebra. For a measurable rectangle $A \times B$, we have that $\pi_x^{-1}(A \times B)$ is $B$ if $x \in A$ and $\varnothing$ otherwise; in both cases, $\pi_x^{-1}(A \times B) \in \mathcal{B}$. So $\mathcal{R}$ contains all the measurable rectangles; as $\mathcal{A} \boxdot \mathcal{B}$ is generated by the rectangles, $\mathcal{R} = \mathcal{A} \boxdot \mathcal{B}$. Hence $E_x = \pi_x^{-1}(E) \in \mathcal{B}$ for all $E \in \mathcal{A} \boxdot \mathcal{B}$. The argument for $E^y$ is entirely analog. $\square$

While we have just shown that the sections are measurable, in general we will want to work in a $\sigma$-algebra bigger than $\mathcal{A} \boxdot \mathcal{B}$ (mostly because having a complete measure space is useful). For such larger $\sigma$-algebras, it is not true in general that a section of a measurable set is always measurable. We will address this in Proposition 2.7.10.

A related notion to that of section is that of **projection**, where the projection onto the second coordinate is $\pi_2 : X \times Y \to Y$ given by $\pi_2(x, y) = y$ and $\pi_1$ is defined analogously. While the two notions are related via $E_x = \pi_2(E \cap (\{x\} \times Y))$, they are very different. Projections were involved in a famous mistake by Lebesgue; the mistake was to assume that the projection of a Borel set in $\mathbb{R}^2$ onto its first coordinate, say, is Borel. This was noted by Suslin, who was a student of Lusin. Suslin began the study of the continuous images of Borel measurable sets on a Polish space, and called these the **analytic sets**; this study started the prolific field of Descriptive Set Theory.

It is not hard to check that for separable metric spaces $X$ and $Y$, $\mathcal{B}(X \times Y) = \mathcal{B}(X) \boxdot \mathcal{B}(Y)$ (Exercise 2.7.4).

Given two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ we want to define the product measure. We naturally expect that for a measurable rectangle $A \times B$ we will have $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$. The task is to construct a measure out of this idea. And the trick we will use is the one we already used to construct Lebesgue measure, which was to construct an outer measure based on the notion of length of an interval.

**2.7.3. Definition.**

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. The* ***product outer measure*** $(\mu \times \nu)^* : \mathcal{P}(X \times Y) \to [0, \infty]$ *is*

$$(\mu \times \nu)^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j) : E \subset \bigcup_{j=1}^{\infty} (A_j \times B_j),\ A_j \in \mathcal{A},\ B_j \in \mathcal{B} \right\},$$

*where we only consider disjoint unions of rectangles.*

The requirement that the unions are disjoint is not a big deal, in light of Lemma 2.7.4 below; in fact, given any countable union of rectangles one can put below it a pairwise disjoint countable union of rectangles such that the series of the products of the measures of the rectangles is smaller (Exercise 2.7.7). A simple observation we will need is that intersections of rectangles are again rectangles, and complements of rectangles are disjoint unions of rectangles. Indeed,

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D), \tag{2.33}$$

and

$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c). \tag{2.34}$$

It is instructive to draw these two identities in the plane, and to see that there are several ways to write $(A \times B)^c$ as a disjoint union of rectangles. The important point is that all set operations on rectangles can be expressed as disjoint unions of rectangles:

***2.7.4. Lemma.***

*Let* $\mathcal{R} = \left\{ \bigcup_{k=1}^{m} A_k \times B_k : \ m \in \mathbb{N}, \ A_k \in \mathcal{A}, \ B_k \in \mathcal{B} \right\}$*, and* $\mathcal{S}$ *the set of all finite unions, intersections, differences, and complements of measurable rectangles. Then* $\mathcal{S} = \mathcal{R}$*.*

*Also, a finite intersection of countable unions of rectangles is a countable union of rectangles, and any countable union of rectangles can be written as a countable disjoint union of rectangles, of the form*

$$\bigcup_j C_j \times D_j, \qquad \text{with } \{C_j\} \text{ pairwise disjoint.}$$

*Proof.* Noting that any combination of set operations can be written in terms of unions and intersections, the equalities (2.33) and (2.34) allow us to write any combination of set operations on a finite number of rectangles as a union of rectangles. Hence $\mathcal{S} = \mathcal{R}$.

For the second assertion, if $\{E_k\}$ and $\{F_j\}$ are countable families of rectangles, then

$$\left(\bigcup_k E_k\right) \cap \left(\bigcup_j F_j\right) = \bigcup_{k,j} (E_k \cap F_j).$$

Now suppose that $E = \bigcup_n A_n \times B_n$ is a countable union of rectangles. Consider all the subsets

$$A_n \setminus \bigcup_{k \neq n} A_k, \qquad \text{and} \qquad A_k \cap A_\ell \ \text{ for } k \neq \ell$$

and (after removing possible repetitions and empty sets) relabel them $\{C_j\}$, which are pairwise disjoint by construction. Similarly, relabel the subsets

$$B_n \setminus \bigcup_{k \neq n} B_k, \qquad \text{and} \qquad B_k \cap B_\ell \ \text{ for } k \neq \ell$$

as $D'_1, D'_2, \ldots$ Let $I_j = \{s : \ C_j \times D'_s \subset E\}$ and form

$$D_j = \bigcup_{s \in I_j} D'_s.$$

We claim that

$$E = \bigcup_j C_j \times D_j.$$

Indeed, if $(c, d) \in C_j \times D_j$ then by definition of $D_j$ there exists $s$ with $(c, d) \in C_j \times D'_s \subset E$. Conversely, if $(a, b) \in A_k \times B_k$, then as $a \in A_k$ and $b \in B_k$ there exist $j$ and $s$ such that $a \in C_j$ and $s$ with $b \in D'_s$. Hence $b \in D_j$, so $(a, b) \in C_j \times D_j$. $\square$

It is worth mentioning that a countable intersection of (finite) unions of rectangles is not necessarily a countable union of rectangles, so $\mathcal{A} \boxdot \mathcal{B}$ contains more sets than just countable unions of rectangles. For instance in $\mathbb{R} \times \mathbb{R}$, with the Lebesgue $\sigma$-algebra in both cases,

$$\bigcap_n \bigcup_{k=1}^n \left[\frac{k-1}{n}, \frac{k}{n}\right]^2 = \{(x,x) :\ 0 \leq x \leq 1\},$$

which is not a countable union of rectangles (Exercise 2.7.6).

**2.7.5. Proposition.**

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. Then $(\mu \times \nu)^*$ is an outer measure and $(\mu \times \nu)^*(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$.*

*Proof.* We have $(\mu \times \nu)^*(\varnothing) = 0$ since $\varnothing \times \varnothing$ appears as one of the covers. The definition via an infimum guarantees that $E \subset F$ implies $(\mu \times \nu)^*(E) \leq (\mu \times \nu)^*(F)$. Now, for the $\sigma$-subadditivity, let $\{E_k\}$ be a countable family with $E_k \subset X \times Y$ for all $k$. If $(\mu \times \nu)^*(E_k) = \infty$ for some $k$, the subadditivity holds trivially. So we may assume that $(\mu \times \nu)^*(E_k) < \infty$ for all $k$. Fix $\varepsilon > 0$ and for each $k$ choose a countable family $\{A_{k,j} \times B_{k,j}\}_j$ such that

$$E_k \subset \bigcup_j (A_{k,j} \times B_{k,j}), \qquad \sum_j \mu(A_{k,j})\nu(B_{k,j}) \leq (\mu \times \nu)^*(E_k) + \frac{\varepsilon}{2^k}.$$

Then $\{A_{k,j} \times B_{k,j}\}_{k,j}$ covers $\bigcup\limits_k E_k$ and (using the definition of $(\mu \times \nu)^*$ for the first inequality)

$$\begin{aligned}(\mu \times \nu)^*\Big(\bigcup_n E_k\Big) &\leq \sum_{k,j} \mu(A_{k,j})\nu(B_{k,j}) \leq \sum_k (\mu \times \nu)^*(E_k) + \frac{\varepsilon}{2^k} \\ &= \varepsilon + \sum_k (\mu \times \nu)^*(E_k).\end{aligned}$$

As this can be done for any $\varepsilon > 0$, the $\sigma$-subadditivity is proven.

Surprisingly, it is not trivial to show that $(\mu \times \nu)^*(A \times B) = \mu(A)\nu(B)$; maybe the surprise should not be that big when we remember that in Proposition 2.3.12 it was not automatic that $m^*([a,b]) = b - a$. If $A \in \mathcal{A}$, $B \in \mathcal{B}$, we have by definition that $(\mu \times \nu)^*(A \times B) \leq \mu(A)\nu(B)$. When $A \times B \subset \bigcup\limits_k R_k$ for pairwise disjoint rectangles $R_k = A_k \times B_k$, we proceed as follows. We have

$$A \times B = \bigcup_k (A_k \cap A) \times (B_k \times B).$$

Because the union is disjoint,

$$1_A(x)\,1_B(y) = 1_{A\times B}(x,y) = \sum_k 1_{(A_k\cap A)\times(B_k\cap B)}(x,y)$$
$$= \sum_k 1_{A_k\cap A}(x) 1_{B_k\cap B}(y).$$

Using Monotone Convergence twice,

$$\begin{aligned}
\mu(A)\nu(B) &= \int_X \Big(\int_Y 1_A(x)1_B(y)\,d\nu(y)\Big)\,d\mu(x)\\
&= \int_X \Big(\int_Y \sum_k 1_{A_k\cap A}(x)1_{B_k\cap B}(y)\,d\nu(y)\Big)\,d\mu(x)\\
&= \int_X \sum_{k,j} \int_Y 1_{A_k\cap A}(x)1_{B_k\cap B}(y)\,d\nu(y)\,d\mu(x)\\
&= \int_X \sum_k 1_{A_k\cap A}(x)\nu(B_k\cap B)\,d\mu(x)\\
&= \sum_k \int_X 1_{A_k\cap A}(x)\nu(B_k\cap B)\,d\mu(x) = \sum_k \mu(A_k\cap A)\nu(B_k\cap B)\\
&\le \sum_k \mu(A_k)\nu(B_k).
\end{aligned}$$

As this can be done for any pairwise disjoint countable family of rectangles that covers $A\times B$, this implies that $\mu(A)\nu(B) \le (\mu\times\nu)^*(A\times B)$, and thus the equality $(\mu\times\nu)^*(A\times B) = \mu(A)\nu(B)$ follows. $\square$

With an outer measure at our disposal Carathéodory's Theorem (2.3.14) provides us with a $\sigma$-algebra $\mathcal{M}(X\times Y)$ where $(\mu\times\nu)^*$ becomes a measure $\mu\times\nu$. Recall from Lemma 2.3.20 that $\mu\times\nu$ is complete. First order of business is to check that $\mathcal{M}(X\times Y)$ is big enough.

***2.7.6. Proposition.***

*Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be measure spaces. Then $\mathcal{A}\boxdot\mathcal{B}\subset\mathcal{M}(X\times Y)$.*

*Proof.* It is enough to show that rectangles are in $\mathcal{M}(X\times Y)$, since the latter is a $\sigma$-algebra and the former generate $\mathcal{A}\boxdot\mathcal{B}$. Let $A\in\mathcal{A}$, $B\in\mathcal{B}$, $S\subset X\times Y$, and $\varepsilon>0$. By definition of the outer measure there exists a disjoint countable family of measurable rectangles $\{A_k \times B_k\}$ such that $S \subset \bigcup_k (A_k \times B_k)$ and

$$(\mu \times \nu)^*(S) + \varepsilon \geq \sum_k \mu(A_k)\nu(B_k).$$

We can write, using (2.34),

$$\begin{aligned}(A \times B)^c \cap \bigcup_k (A_k \times B_k) = \bigcup_k \big((A_k \cap A^c) \times (B_k \cap B)\big) \\ \cup \bigcup_k \big((A_k \cap A) \times (B_k \cap B^c)\big) \\ \cup \bigcup_k \big((A_k \cap A^c) \times (B_k \cap B^c)\big).\end{aligned}$$

Then, using the definition of the outer measure in the middle inequality,

$$\begin{aligned}(\mu \times \nu)^*(S) + \varepsilon &\geq \sum_k \mu(A_k)\nu(B_k) \\ &= \sum_k \big[\mu(A_k \cap A) + \mu(A_k \cap A^c)\big]\big[\nu(B_k \cap B) + \nu(B_k \cap B^c)\big] \\ &= \sum_k \mu(A_k \cap A)\nu(B_k \cap B) + \sum_k \mu(A_k \cap A)\nu(B_k \cap B^c) \\ &\qquad + \mu(A_k \cap A^c)\nu(B_k \cap B) + \mu(A_k \cap A^c)\nu(B_k \cap B^c) \\ &\geq (\mu \times \nu)^*\Big((A \times B) \cap \bigcup_k (A_k \times B_k)\Big) \\ &\qquad + (\mu \times \nu)^*\Big((A \times B)^c \cap \bigcup_k (A_k \times B_k)\Big) \\ &\geq (\mu \times \nu)^*((A \times B) \cap S) + (\mu \times \nu)^*((A \times B)^c \cap S).\end{aligned}$$

As this holds for all $\varepsilon > 0$, we get

$$(\mu \times \nu)^*(S) = (\mu \times \nu)^*((A \times B) \cap S) + (\mu \times \nu)^*((A \times B)^c \cap S),$$

since the reverse inequality always holds for an outer measure. Then $A \times B \in \mathcal{M}(X \times Y)$. □

The inclusion $\mathcal{A} \boxdot \mathcal{B} \subset \mathcal{M}(X \times Y)$ is usually proper. For example, if $E' \subset [0,1]$ is not Lebesgue measurable, the set $E = E' \times \{0\} \subset \mathbb{R} \times \mathbb{R}$ can be covered by rectangles of the form $[0,1] \times (-\varepsilon, \varepsilon)$ for any $\varepsilon > 0$. This shows that $(m \times m)^*(E) \leq 2\varepsilon$ for all $\varepsilon > 0$ and so $(m \times m)^*(E) = 0$. We learned in Lemma 2.3.20 that such $E$ is measurable. That is, $E \in \mathcal{M}(\mathbb{R} \times \mathbb{R})$ but $E \notin \mathcal{M}(\mathbb{R}) \boxdot \mathcal{M}(\mathbb{R})$; for otherwise $E' = E^0 \in \mathcal{M}(\mathbb{R})$ by Proposition 2.7.2.

A useful consequence of Lemma 2.7.4 is the following.

**2.7.7. Lemma.**

*Let $E \in \mathcal{M}(X \times Y)$ with $(\mu \times \nu)(E) < \infty$. Then there exists a sequence $\{E_k\} \subset \mathcal{A} \boxdot \mathcal{B}$, where each $E_k$ is a countable union of measurable rectangles, such that $E_{k+1} \subset E_k$ for all $k$, $(\mu \times \nu)(E_1) < \infty$,*

$$E \subset \bigcap_k E_k, \qquad \text{and} \qquad (\mu \times \nu)(E) = \lim_{k\to\infty} (\mu \times \nu)(E_k).$$

*As a consequence, every $E \in \mathcal{M}(X \times Y)$ is of the form $E = F \cup G$, with $F \in \mathcal{A} \boxdot \mathcal{B}$ and $G$ a nullset.*

*Proof.* By definition of the outer measure, there exists a union of measurable rectangles $E_1$ such that $E \subset E_1$ and $(\mu\times\nu)(E_1) < (\mu\times\nu)(E)+1$. Inductively, given countable unions of measurable rectangles $E_1 \supset \cdots \supset E_k \supset E$, again by the definition of the outer measure there exists a countable union $F_{k+1}$ of measurable rectangles such that $E \subset F_{k+1}$ and $(\mu\times\nu)(F_{k+1}) < (\mu\times\nu)(E) + \frac{1}{k+1}$. Then we put $E_{k+1} = F_{k+1} \cap E_k$. We have

$$(\mu \times \nu)(E) \le (\mu \times \nu)(E_k) \le (\mu \times \nu)(F_k) < (\mu \times \nu)(E) + \frac{1}{k},$$

showing that $\lim_{k\to\infty}(\mu \times \nu)(E_k) = (\mu \times \nu)(E)$. Using continuity of the measure (Proposition 2.3.8), that applies because $E_1$ has finite measure,

$$(\mu \times \nu)\Big(\bigcap_k E_k\Big) = \lim_{k\to\infty} (\mu \times \nu)(E_k) = (\mu \times \nu)(E).$$

Now we want to use the above to write $E$ as $F \cup G$ with $F \in \mathcal{A} \boxdot \mathcal{B}$ and $G$ a nullset. The problem we have is that the inclusion $E \subset \bigcap_k E_k$ is the wrong way around, as we need an element of $\mathcal{A}\boxdot\mathcal{B}$ *inside* $E$. What we do is to apply the above to $E_1 \setminus E$; we get a decreasing sequence $\{F_k\}$ of countable unions of rectangles with $E_1 \setminus E \subset \bigcap_k F_k$, and $\lim_k(\mu\times\nu)(F_k) = (\mu\times\nu)(E_1 \setminus E)$. If we replace each $F_k$ with $E_1 \cap F_k$ (which is still a countable union of rectangles by Lemma 2.7.4), both the inclusion and the limit still hold, since $(\mu\times\nu)(F_k) \ge (\mu \times \nu)(E_1 \cap F_k) \ge (\mu \times \nu)(E)$. Now

$$E = E_1 \setminus (E_1 \setminus E) \supset E_1 \setminus \bigcap_k F_k = \bigcup_k (E_1 \setminus F_k),$$

and by continuity of the measure

$$\begin{aligned}(\mu \times \nu)\Big(\bigcup_k (E_1 \setminus F_k)\Big) &= \lim_k (\mu \times \nu)(E_1 \setminus F_k) \\ &= (\mu \times \nu)(E_1) - \lim_k (\mu \times \nu)(F_k) \\ &= (\mu \times \nu)(E_1) - (\mu \times \nu)(E_1 \setminus E) \\ &= (\mu \times \nu)(E).\end{aligned}$$

So $F = \bigcup_k (E_1 \setminus F_k) \in \mathcal{A} \boxdot \mathcal{B}$ with $F \subset E$ and $(\mu \times \nu)(F) = (\mu \times \nu)(E)$. Then $E = F \cup G$, with $F \in \mathcal{A} \boxdot \mathcal{B}$ and $G$ a nullset. $\square$

**2.7.1. Lebesgue Measure in $\mathbb{R}^n$.** A canonical place where to construct the product measure is $\mathbb{R}^2$. We apply the above to $(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$ to obtain a product measure on $\mathbb{R} \times \mathbb{R}$. Because $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, by Proposition 2.7.6 the sets of the form $V \times W$ with $V, W \in \mathcal{B}(\mathbb{R})$ are in $\mathcal{M}(\mathbb{R} \times \mathbb{R})$. Thus $\mathcal{B}(\mathbb{R}) \boxdot \mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R} \times \mathbb{R})$. Since in $\mathbb{R}^2$ every open set is a countable union of rectangles, we get that $\mathcal{B}(\mathbb{R}) \boxdot \mathcal{B}(\mathbb{R})$ is a $\sigma$-algebra that contains all open sets and thus $\mathcal{B}(\mathbb{R} \times \mathbb{R}) \subset \mathcal{B}(\mathbb{R}) \boxdot \mathcal{B}(\mathbb{R})$. As the reverse inclusion always occurs, we get $\mathcal{B}(\mathbb{R} \times \mathbb{R}) = \mathcal{B}(\mathbb{R}) \boxdot \mathcal{B}(\mathbb{R})$. More generally, $\mathcal{B}(X) \boxdot \mathcal{B}(Y) = \mathcal{B}(X \times Y)$ whenever $X$ and $Y$ are separable metric spaces (Exercise 2.7.4). The inclusion $\mathcal{B}(X) \boxdot \mathcal{B}(Y) \subset \mathcal{B}(X \times Y)$ can be proper for other topological spaces.

With a bit of work, Lemma 2.7.7 allows us to show that $m \times m$ is outer regular (Exercise 2.7.8).

As $(m \times m)(E \times F) = m(E)m(F)$ for any $E, F \in \mathcal{M}(\mathbb{R})$, the measure $m \times m$ respects the natural measure of rectangles. From $E \times F + (x, y) = (E + x) \times (F + y)$ we get that $m \times m$ is translation invariant on rectangles. Writing open sets as countable unions of rectangles we get that $m \times m$ is translation invariant on open sets. And by the outer regularity, $m \times m$ is translation invariant on all measurable sets.

Now for $n \geq 3$, we can iterate the above doing first $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$, and so on. All the same arguments apply.

As is customary, instead of writing $m \times m$, or $m \times m \times m$, or any higher-dimensional version, we simply write $m$ for the Lebesgue measure on $\mathbb{R}^n$. Context should always make it clear which $m$ we are talking about. There is no commonly accepted standard on how to write Lebesgue measure. All of $m(E)$, $\lambda(E)$, and $|E|$ a very common.

**2.7.2. Measurable Functions and Integration.** Before discussing integration we need to consider the measurable functions.

**2.7.8. Definition.**

*Given $f : X \times Y \to \mathbb{C}$, $x \in X$, $y \in Y$, the **$x$-section** of $f$ is the function $f_x : Y \to \mathbb{C}$ given by $f_x(y) = f(x, y)$, and the **$y$-section** is $f^y : X \to \mathbb{C}$ where $f^y(x) = f(x, y)$.*

The sections of $f$ are relevant because we need them to be measurable for the iterated integrals in (2.35) below to even make sense. A straightforward computation shows that $f_x$ and $f^y$ are measurable if $f$ is $\mathcal{A} \boxdot \mathcal{B}$-measurable (Exercise 2.7.9). The same is not true if $f$ is $\mathcal{M}(X \times Y)$-measurable (Exercise 2.7.11), but if $\mu$ and $\nu$ are complete, then $f_x$ is measurable $\nu$-a.e. and $f^y$ is measurable $\mu$-a.e. (Exercise 2.7.10).

In the following we will have several $\sigma$-algebras around; for starters, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{A} \boxdot \mathcal{B}$, and also $\mathcal{M}(X \times Y)$. When we talk measurability, it should be clear by context to which of them we refer. The exception is for functions $f : X \times Y \to \mathbb{C}$. For such functions it is equally natural to consider measurability with respect to $\mathcal{A} \boxdot \mathcal{B}$ and with respect to $\mathcal{M}(X \times Y)$. Our choice will always be the latter, unless explicitly mentioned. This has the advantage that it makes us work in a complete measure space.

When integrating on a product space, we want to be able to integrate in terms of the individual measures. For this to make sense, the iterated integrals should not depend on the order. That is, we expect

$$\int_{X\times Y} f\,d(\mu \times \nu) = \int_X \int_Y f(x,y)\,d\mu(y)\,d\mu(x) = \int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y). \tag{2.35}$$

While these equalities are regularly used in several-variable calculus courses, there is no reason a priori why they would hold in general; the point here is precisely to get conditions that guarantee that (2.35) holds; Fubini's Theorem (Theorems 2.7.12 and 2.7.16) does precisely that.

We will assume from now on that our spaces are complete. This is not a big deal as we can complete the measure if needed (Proposition 2.3.21). The following two results contain the key technical ingredients for the proof of Fubini's Theorem. Lemma 2.7.9 is just a part of the proof of Proposition 2.7.10 that we isolate for clarity—hence the weird hypotheses.

***2.7.9. Lemma.***

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces and $G \in \mathcal{M}(X \times Y)$ a nullset. Suppose that the functions $x \longmapsto \nu(E_x)$ are measurable and that $\int_X \nu(E_x)\,d\mu = (\mu \times \nu)(E)$ for all countable intersections of countable unions of rectangles $E \in \mathcal{A} \boxdot \mathcal{B}$. If furthermore $\nu$ is complete, then $G_x$ is measurable and a nullset $\mu$-a.e.*

*Proof.* Assuming $(\mu\times\nu)(G) = 0$, let $\{E_k\}$ as in Lemma 2.7.7. Put $F = \bigcap_k E_k$. From $E_k \in \mathcal{A} \boxdot \mathcal{B}$, we have that $(E_k)_x$ is measurable by Proposition 2.7.2. Then $F_x = \bigcap_k (E_k)_x$ is $\nu$-measurable for all $x$.

Since $(\mu \times \nu)(E_1) < \infty$, by definition of the product measure there exist rectangles $A_k, B_k$ such that $E_1 = \bigcup_k A_k \times B_k$ and $\sum_k \mu(A_k)\nu(B_k) < 1 + (\mu \times$

$\nu)(E_1)$. Note that

$$(E_1)_x = \bigcup_k (A_k \times B_k)_x = \bigcup_{k:\ x \in A_k} B_k,$$

since $(A_k \times B_k)_x = \varnothing$ when $x \notin A_k$.

Then, using Corollary 2.5.8,

$$\begin{aligned}\int_X \nu((E_1)_x)\, d\mu(x) &= \int_X \nu\Big(\bigcup_{k:\ x\in A_k} B_k\Big)\, d\mu(x) \leq \int_X \sum_k \nu(B_k)\, 1_{A_k}(x)\, d\mu(x)\\ &= \sum_k \int_X \nu(B_k)\, 1_{A_k}(x)\, d\mu(x) = \sum_k \mu(A_k)\nu(B_k)\\ &< 1 + (\mu\times\nu)(E_1) < \infty.\end{aligned}$$

As $\nu((E_k)_x) \leq \nu((E_1)_x)$, by Dominated Convergence and continuity of the measure

$$\int_X \nu(F_x)\, d\mu(x) = \lim_{k\to\infty} \int_X \nu((E_k)_x)\, d\mu(x) = \lim_{k\to\infty} (\mu\times\nu)(E_k) = 0.$$

Hence $\nu(F_x) = 0$ a.e. $(\mu)$. For those $x$ such that $\nu(F_x) = 0$, as $G_x \subset F_x$ we get that $G_x$ is measurable and $\nu(G_x) = 0$ by the completeness of $(Y, \mathcal{B}, \nu)$. So $G_x$ is measurable and $\nu(G_x) = 0$ a.e. $(\mu)$. □

***2.7.10. Proposition.***

*Let $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ be complete measure spaces and $E \in \mathcal{M}(X \times Y)$ such that $(\mu\times\nu)(E) < \infty$. Then $E_x$ is $\nu$-measurable for $\mu$-almost all $x \in X$ and $E^y$ is $\mu$-measurable for $\nu$-almost all $y \in Y$; and*

$$(\mu\times\nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y) \tag{2.36}$$

*Proof.* We only need to work with $E_x$ and prove the first equality, as the other will have an analog proof by exchanging the roles of $x$ and $y$ (and of $\mu$ and $\nu$). We will go through different cases of what $E$ could be.

Assume first that $E = A \times B$, with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Then $E_x = B$ when $x \in A$ and $\varnothing$ when $x \notin A$, so $\nu$-measurable for all $x$. As $\nu(E_x) = 1_A(x)\,\nu(B)$, it is $\mu$-measurable as a function of $x$ since $A \in \mathcal{A}$. And (with Proposition 2.7.5 for the second to last equality)

$$\int_X \nu(E_x)\, d\mu(x) = \nu(B) \int_X 1_A\, d\mu = \mu(A)\nu(B) = (\mu\times\nu)(A\times B) = (\mu\times\nu)(E).$$

Next suppose that $E = \bigcup_k A_k \times B_k$ a countable union of rectangles, with $A_k \in \mathcal{A}$, $B_k \in \mathcal{B}$ for all $k$. Write $E_k = A_k \times B_k$. By Lemma 2.7.4 we may assume without loss of generality that the union is disjoint and that the $\{B_k\}$ are pairwise disjoint. The set $E_x = \bigcup_k (E_k)_x = \bigcup_{x \in A_k} B_k$ is measurable. Because the $B_k$ are pairwise disjoint, $\nu(E_x) = \sum_k \nu((E_k)_x)$ so $\mu$-measurable, and then (using Corollary 2.5.8 and the previous case)

$$\int_X \nu(E_x)\, d\mu(x) = \sum_k \int_X \nu((E_k)_x)\, d\mu(x) = \sum_k (\mu \times \nu)(E_k) = (\mu \times \nu)(E).$$

We now move to the case where $E = \bigcap_k E_k$ and each $E_k$ is a countable union of rectangles with $E_{k+1} \subset E_k$. By replacing $E_k$ with $E_k \cap E$ we may assume without loss of generality that $(\mu \times \nu)(E_k) < \infty$ for all $k$. By the previous paragraph we have that $(E_k)_x$ is $\nu$ measurable for all $x$. Also,

$$\int_X \nu((E_1)_x)\, d\mu(x) = (\mu \times \nu)(E_1) < \infty,$$

and so $\nu((E_1)_x) < \infty$ a.e.$(\mu)$. As $E_x = \bigcap_k (E_k)_x$ we have that $\nu(E_x) = \lim_k \nu((E_k)_x)$ by continuity of the measure so $x \longmapsto \nu(E_x)$ is measurable. By Dominated Convergence and continuity of the measure

$$\int_X \nu(E_x)\, d\mu(x) = \lim_{k\to\infty} \int_X \nu((E_k)_x)\, d\mu(x) = \lim_{k\to\infty} (\mu \times \nu)(E_k) = (\mu \times \nu)(E).$$

We can now move into the general case, where $E \in \mathcal{M}(X \times Y)$ with finite measure. Let $\{E_k\}$ be as in Lemma 2.7.7. Write $F = \bigcap_k E_k$ (attention! This is not the $F$ from Lemma 2.7.7). By the case above, $F_x$ is measurable for all $x$, and so is $x \longmapsto \nu(F_x)$. As $(\mu \times \nu)(F \setminus E) = 0$, by Lemma 2.7.9 the section $(F \setminus E)_x$ is $\nu$-measurable and $\nu((F \setminus E)_x) = 0$ a.e. $(\mu)$. Then $E_x = F_x \setminus (F \setminus E)_x$ is $\nu$-measurable a.e. $(\mu)$. Also by the previous paragraph in the case of $F$ and by Lemma 2.7.9 in the case of $F \setminus E$ (as the latter is a nullset),

$$(\mu \times \nu)(F) = \int_X \nu(F_x)\, d\mu(x), \qquad (\mu \times \nu)(F \setminus E) = \int_X \nu(F_x \setminus E_x)\, d\mu(x).$$

From this,

$$\begin{aligned}
(\mu \times \nu)(E) = (\mu \times \nu)(F \setminus (F \setminus E)) &= (\mu \times \nu)(F) - (\mu \times \nu)(F \setminus E) \\
&= \int_X \big[\nu(F_x) - \nu(F_x \setminus E_x)\big]\, d\mu(x) \\
&= \int_X \nu(E_x)\, d\mu(x).
\end{aligned}$$

□

The name "Fubini's Theorem" is applied to a family of results that guarantee that the double integral agrees with the iterated integrals. Here we split the results in two due to the fact that to go from an integrable $f$ to the iterated integrals less hypotheses are needed; concretely, $\sigma$-finiteness is required to go from the iterated integrals to the double integral but not the other way around.

*2.7.11. Remark.* The completeness requirement was needed only at an obscure place in the proof of Proposition 2.7.10. So, is it really necessary in Proposition 2.7.10 and, as a consequence, in Theorems 2.7.12 and 2.7.16? To confuse things even more, we showed in Proposition 2.7.2 that sections were measurable. The key observation is that in Proposition 2.7.2 we were working on $\mathcal{A}\boxdot\mathcal{B}$, and now we are working on the bigger $\sigma$-algebra $\mathcal{M}(X\times Y)$. Completeness is not needed in Theorems 2.7.12 and 2.7.16 when $f$ is required to be $\mathcal{A}\boxdot\mathcal{B}$-measurable instead.

**2.7.12. Theorem. (Fubini, Part 1)**

*Let $(X,\mathcal{A},\mu)$, $(Y,\mathcal{B},\nu)$ be complete measure spaces and $f: X\times Y\to\mathbb{C}$ be $(\mu\times\nu)$-integrable. Then $f_x$ is $\nu$-integrable a.e. $(\mu)$, $f^y$ is $\mu$-integrable a.e. $(\nu)$, the functions $x\longmapsto\int_Y f_x\,d\nu$ and $y\longmapsto\int_X f^y\,d\mu$ are respectively $\mu$-integrable and $\nu$-integrable, and*

$$\begin{aligned}\int_{X\times Y} f\,d(\mu\times\nu) &= \int_X\int_Y f(x,y)\,d\nu(y)\,d\mu(x)\\ &= \int_Y\int_X f(x,y)\,d\mu(x)\,d\nu(y)\end{aligned}\tag{2.37}$$

*Proof.* When $f=1_E$ for $E\in\mathcal{M}(X\times Y)$, the result was established in Proposition 2.7.10 (note that $(1_E)_x = 1_{E_x}$). As a simple function $s=\sum_k a_k\,1_{E_k}$ is integrable if and only if each $1_{E_k}$ is, the theorem follows for simple functions by linearity of the integral and the fact that $s_x=\sum_k a_k\,(1_{E_k})_x$, and similarly with $y$.

If now $f\geq 0$, by Theorem 2.4.13 there exists a sequence $\{s_n\}$ of simple functions with $0\leq s_n\leq s_{n+1}\nearrow f$. Because $f$ is integrable, so is each $s_n$. Then the previous paragraph applies to $s_n$. As $f_x=\lim_n(s_n)_x$ it is measurable. From the previous paragraph we have

$$\int_{X\times Y} s_n\,d(\mu\times\nu) = \int_X\int_Y s_n(x,y)\,d\nu(y)\,d\mu(x) = \int_Y\int_X s_n(x,y)\,d\mu(x)\,d\nu(y).$$

Using Monotone Convergence, we obtain (2.37). As the leftmost integral is finite by hypothesis, the other two are. In particular the two inner integrals, which are non-negative, are integrable a.e. as stated.

When $f$ is real and integrable, so are $f^+$ and $f^-$. Then we obtain (2.37) for $f^+$ and $f^-$, and thus for $f$ by linearity. Similarly, when $f$ is complex valued and integrable, $\operatorname{Re} f$ and $\operatorname{Im} f$ are real valued and integrable, and so again we obtain (2.37) by linearity. □

The second version of Fubini involves the new concept of $\sigma$-finiteness of a measure space.

**2.7.13. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space. We say that $\mu$ is **$\sigma$-finite** if there exist $X_1, X_2, \ldots \in \mathcal{A}$ such that $X = \bigcup_n X_n$ and $\mu(X_n) < \infty$ for all $n$.*

Canonical examples of $\sigma$-finite measures are Lebesgue measure on the real line, and the counting measure on any countable set. An easy example of a non-$\sigma$-finite measure would be the counting measure on an uncountable set. We remark that there is no requirement on the $X_n$ other than being finite-measure, so it is indistinct to require that the $\{X_n\}$ are pairwise disjoint, or that $X_n \subset X_{n+1}$ for all $n$ (Exercise 2.7.12).

Here is a result that will be needed in the future.

***2.7.14. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space. The following statements are equivalent:*

*(i) $\mu$ is $\sigma$-finite;*

*(ii) there exists integrable $f$ such that $0 < f(x) < 1$ for all $x \in X$.*

*Proof.* If $\mu$ is $\sigma$-finite, we can write $X = \bigcup_n X_n$, pairwise disjoint, with $\mu(X_n) < \infty$ for all $n$. Define

$$f = \sum_n \frac{1}{2^n(1 + \mu(X_n))} 1_{X_n}.$$

Then, by definition, $0 < f(x) < 1$ for all $x$, and by Corollary 2.5.8

$$\int_X f \, d\mu = \sum_n \frac{\mu(X_n)}{2^n(1 + \mu(X_n))} < 1.$$

Conversely, if there exists $f$ integrable with $0 < f(x) < 1$ for all $x$, define

$$X_n = \left\{x : \frac{1}{n+1} \leq f(x) < \frac{1}{n}\right\}.$$

Then $X = \bigcup_n X_n$, pairwise disjoint, and for each $n$

$$\mu(X_n) = (n+1)\int_{X_n} \frac{1}{n+1}\,d\mu \le (n+1)\int_{X_n} f\,d\mu < \infty. \qquad \square$$

In the second Fubini theorem below, the assumption that the iterated integral

$$\int_X \int_Y |f(x,y)|\,d\nu(y)\,d\mu(x)$$

exists implies the assumption that $y \longmapsto |f(x,y)|$ is $\nu$-measurable. And we will need the following:

**2.7.15. Lemma.**

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces and let $f : X \times Y \to \mathbb{C}$ be $(\mu \times \nu)$-measurable. Then the functions*

$$x \longmapsto \int_Y f(x,y)\,d\nu(y) \qquad \text{and} \qquad y \longmapsto \int_X f(x,y)\,d\mu(y)$$

*are measurable.*

*Proof.* As usual, we only need to do one of the cases. Assume first that $f = 1_E$ for some measurable $E$. Then we proved in Proposition 2.7.10 that $x \longmapsto \nu(E_x) = \int_y 1_E(x,y)\,d\nu(y)$ is measurable. If $f$ is non-negative and simple, then $x \longmapsto \int_Y f\,d\nu$ is a linear combination of measurable functions and hence measurable. For non-negative measurable $f$, we can write it as a monotone limit of simple functions, and then $x \longmapsto \int_Y f\,d\nu$ is a limit of measurable functions. For arbitrary $f$ we can write it as a linear combination of non-negative functions and the result follows. $\square$

**2.7.16. Theorem. (Fubini, Part 2)**

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be $\sigma$-finite and complete. Let $f : X \times Y \to \mathbb{C}$ be $(\mu \times \nu)$-measurable. If at least one of*

$$\int_X \int_Y |f(x,y)|\,d\nu(y)\,d\mu(x),\ \int_Y \int_X |f(x,y)|\,d\mu(x)\,d\nu(y),\ \int_{X\times Y} |f|\,d(\mu\times\nu)$$

*is finite, then $f$ is integrable, and*

$$\begin{aligned}\int_{X\times Y} f\,d(\mu\times\nu) &= \int_X \int_Y f(x,y)\,d\nu(y)\,d\mu(x) \\ &= \int_Y \int_X f(x,y)\,d\mu(x)\,d\nu(y).\end{aligned} \tag{2.38}$$

*Proof.* The case where $f$ is integrable is covered by Theorem 2.7.12. By Exercise 2.7.10 we know that $f_x$ and $f^y$ are measurable a.e. We will assume that

$$\int_X \int_Y |f(x,y)|\, d\nu(y)\, d\mu(x) < \infty,$$

as the other case is similar. By Lemma 2.7.15 the iterated integrals make sense for each of $(\operatorname{Re} f)^+$, $(\operatorname{Re} f)^-$, $(\operatorname{Im} f)^+$, $(\operatorname{Im} f)^-$. Hence we may assume without loss of generality that $f \geq 0$.

By Lemma 2.5.13, $\int_Y f(x,y)\, d\nu(y) < \infty$ a.e. $(\mu)$, so $f_x$ is integrable a.e. $(\mu)$. By writing $X = \bigcup_m X_m$, $Y = \bigcup_n Y_n$ as increasing unions of finite-measure sets, we can write $f = \lim_n f\, 1_{X_n} 1_{Y_n}$ as an non-decreasing limit of functions with finite-measure support. And approximating each $f\, 1_{X_n} 1_{Y_n}$ by increasing sequences of simple functions, we can obtain a sequence $\{s_n\}$ of simple functions such that $s_n \nearrow f$ and each $s_n$ is integrable (as a simple function on a finite-measure space is always integrable). So Theorem 2.7.12 applies to each $s_n$, and then using Monotone Convergence first once and then twice we get

$$\begin{aligned}\int_{X\times Y} f\, d(\mu\times\nu) = \lim_n \int_{X\times Y} s_n\, d(\mu\times\nu) &= \lim_n \int_X \int_Y s_n(x,y)\, d\nu(y)\, s\mu(x) \\ &= \int_X \int_Y f(x,y)\, d\nu(y)\, d\mu(x).\end{aligned}$$

Then $f$ is integrable and by Theorem 2.7.12 we obtain (2.38). $\square$

A commonly used consequence of Fubini's theorem is the following. The reason it deserves to be mentioned on its own is the fact that no integrability is mentioned, i.e., (2.39) holds even if the integrals are infinite.

***2.7.17. Corollary. (Tonelli)***

*Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be $\sigma$-finite and complete. Let $f : X \times Y \to [0,\infty]$ be $(\mu\times\nu)$-measurable. Then*

$$\int_{X\times Y} f\, d(\mu\times\nu) = \int_X \int_Y f(x,y)\, d\nu(y)\, d\mu(x) = \int_Y \int_X f(x,y)\, d\mu(x)\, d\nu(y). \tag{2.39}$$

*Proof.* If at least one of the three integrals in (2.39) is finite, then all three are finite and equal by Theorem 2.7.16. Otherwise, all three are infinite and thus equal. $\square$

We addressed the necessity of the completeness hypothesis in Remark 2.7.11.

*2.7.18. Remark* (Counterexamples).

- **$\sigma$-finiteness** Let $X = [0,1]$ with Lebesgue measure, and $Y = [0,1]$ with the counting measure (so all sets are $\nu$-measurable and $\nu$ is not $\sigma$-finite). Let $E = \{(x,x) :\ x \in [0,1]\}$; this set is measurable because
$$E = \bigcap_n \bigcup_{k=1}^n \left[\frac{k-1}{n}, \frac{k}{n}\right]^2. \tag{2.40}$$
We have
$$\int_{[0,1]}\int_{[0,1]} 1_E(x,y)\,d\nu(y)\,d\mu(x) = \int_{[0,1]} \nu(\{x\})\,1_E(x,x)\,d\mu$$
$$= \int_{[0,1]} 1_E(x,x)\,d\mu = \mu([0,1]) = 1,$$
while
$$\int_{[0,1]}\int_{[0,1]} 1_E(x,y)\,d\mu(x)\,d\nu(y) = \int_{[0,1]} \mu(\{y\})\,1_E(y,y)\,d\nu(y) = 0.$$
So Tonelli can fail when one of the spaces is not $\sigma$-finite. Even worse, $(\mu\times\nu)(E) = \infty$. Indeed, if $(\mu\times\nu)(E) < \infty$ it would mean that $1_E$ was integrable, and in that case Theorem 2.7.12 would imply that the iterated integrals were equal. Hence all three integrals in (2.39) can be distinct in the non-$\sigma$-finite case.
- **Non-integrable functions**. Fubini can fail when $f$ is not integrable. A typical example, on $[0,1]^2 \subset \mathbb{R}^2$ with Lebesgue measure, is
$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}$$
(defined as 0 at the origin, say). Then
$$\int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy = \int_0^1 \left.\frac{-x}{x^2+y^2}\right|_{x=0}^{x=1} dy = -\int_0^1 \frac{1}{y^2+1}\,dy$$
$$= -\arctan 1 = -\frac{\pi}{4},$$
while
$$\int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\,dx = \int_0^1 \left.\frac{y}{x^2+y^2}\right|_{y=0}^{y=1} dx = \int_0^1 \frac{1}{x^2+1}\,dx = \arctan 1 = \frac{\pi}{4}.$$

The different values of the iterated integrals already show that $f$ is not integrable. We can also check this directly, with ease, in the following way. Because of Tonelli's Theorem we can calculate the integral of $|f|$ by using iterated integrals; then
$$\iint_{[0,1]^2} |f| \ge \int_0^1\int_0^x \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\,dx = \int_0^1 \left.\frac{y}{x^2+y^2}\right|_{y=0}^{y=x} dx = \int_0^1 \frac{1}{2x}\,dx = \infty.$$

**2.7.3. Exercises.**

(2.7.1) Let $E, E_k \subset X \times Y$ for all $k \in \mathbb{N}$, and $x \in X$. Show that

$$\Big(\bigcup_k E_k\Big)_x = \bigcup_k (E_k)_x, \qquad \Big(\bigcap_k E_k\Big)_x = \bigcap_k (E_k)_x,$$

and

$$(E^c)_x = (E_x)^c$$

(2.7.2) Write another proof for Proposition 2.7.2 by showing the set

$$\mathcal{S} = \{E \subset X \times Y : \ E_x \in \mathcal{B}, \ E^y \in \mathcal{A} \ \text{ for all } x \in X, \ y \in Y\}.$$

is a $\sigma$-algebra that contains the measurable rectangles.

(2.7.3) Show that every $\sigma$-finite measure $\mu$ is semifinite (see Exercise 2.3.26). Give an example of a semifinite measure that is not $\sigma$-finite.

(2.7.4) Let $X$ and $Y$ be separable metric spaces. Show that $\mathcal{B}(X \times Y) = \mathcal{B}(X) \boxdot \mathcal{B}(Y)$.

(2.7.5) Show that for any $\varepsilon > 0$ there exists $V \subset \mathbb{R}^n$, open, dense, and with $m(V) < \varepsilon$. Then answer the following questions:

(i) Does the measure of an open subset of $\mathbb{R}^n$ agree with the measure of its closure?

(ii) Is the measure of the boundary of every open subset of $\mathbb{R}^n$ zero?

(2.7.6) Show that, in $\mathbb{R} \times \mathbb{R}$, the set

$$\bigcap_n \bigcup_{k=1}^n \left[\frac{k-1}{n}, \frac{k}{n}\right]^2 = \{(x,x) : \ 0 \le x \le 1\} \tag{2.41}$$

is not a countable union of rectangles (measurable or not). Conclude that the set of countable unions of rectangles does not form a $\sigma$-algebra.

(2.7.7) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces and $\{A_k \times B_k\}$ a countable family of rectangles. Use Lemma 2.7.4 to show that there exists a pairwise disjoint countable family of rectangles $\{A_k' \times B_k'\}$ such that

$$\bigcup_k A_k' \times B_k' = \bigcup_k A_k \times B_k$$

and

$$\sum_k \mu(A_k')\nu(B_k') \le \sum_k \mu(A_k)\nu(B_k).$$

(2.7.8) Show that $m \times m$ is outer regular.

(2.7.9) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. Let $f : X \times Y \to \mathbb{C}$ be $\mathcal{A} \boxdot \mathcal{B}$-measurable. Show that the sections $f_x$ and $f^y$ are measurable.

(2.7.10) Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be complete measure spaces, and $f : X \times Y \to \mathbb{C}$ measurable. Show that the sections $f_x$ and $f^y$ are measurable a.e.

(2.7.11) Show an example of an $\mathcal{M}(X \times Y)$-measurable function $f$ such that $f_x$ is not measurable for $x$ in a set of positive measure.

(2.7.12) Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that the following statements are equivalent:

(i) there exist $\{X_n\} \subset \mathcal{A}$, pairwise disjoint, with $\mu(X_n) < \infty$ for all $n$ and $X = \bigcup_n X_n$;

(ii) there exist $\{X_n\} \subset \mathcal{A}$, with $X_n \subset X_{n+1}$ and $\mu(X_n) < \infty$ for all $n$, and with $X = \bigcup_n X_n$.

(2.7.13) Let $f : \mathbb{R}^d \to \mathbb{C}$ be integrable. Show that

$$\int_{\mathbb{R}^d} f(t)\, dm(t) = \int_{\mathbb{R}^d} f(x-t)\, dm(t), \qquad x \in \mathbb{R}^d. \tag{2.42}$$

(2.7.14) Use Fubini's Theorem and the equality

$$\int_0^\infty e^{-xt}\, dt = \frac{1}{x}, \quad x > 0,$$

to show that

$$\int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

Note that $x \longmapsto \frac{\sin x}{x}$ is not integrable.

(2.7.15) Let $f : X \to [0, \infty)$ be measurable. Use the layer-cake representation (2.26) to conclude that

$$\int_X f\, d\mu = \int_0^\infty \mu(\{f \geq t\})\, dt.$$

(2.7.16) Let $(X, \mathcal{A}, \mu)$ be a $\sigma$-finite measure space. Let $f : X \to [0, \infty)$ be measurable.

(i) Show that the graph of $f$ is a $(\mu \times m)$-nullset.

(ii) Show that if

$$B = \{(x, t) \in X \times \mathbb{R} : \ 0 < t < f(x)\}$$

then $(\mu \times m)(B) = \displaystyle\int_X f\, d\mu$.

(iii) Does the above work if we replace the codomain with an arbitrary measure space?

(2.7.17) *(Polar Coordinates in $\mathbb{R}^n$).* Let $S_{n-1}$ be the unit sphere on $\mathbb{R}^n$ (i.e., those $u$ with $\|u\| = 1$). Show that any nonzero $x \in \mathbb{R}^n$ can be written $x = ru$, with $r > 0$ and $u \in S_{n-1}$. Thus $\mathbb{R}^n \setminus \{0\}$ can be seen as the Cartesian product $(0, \infty) \times S_{n-1}$.

Let $m$ be the Lebesgue measure on $\mathbb{R}^n$, and define a measure $\sigma$ on the Borel sets of $S_{n-1}$ by

$$\sigma(A) = n\, m(\tilde{A}),$$

where $\tilde{A} = \{ru :\ 0 < r < 1,\ u \in A\}$. Show that for every Borel $f \geq 0$,

$$\int_{\mathbb{R}^n} f\, dm = \int_0^\infty \int_{S_{n-1}} r^{n-1} f(ru)\, d\sigma(u)\, dr. \tag{2.43}$$

*Hint:* if $A \subset S_{n-1}$ is open and $0 < r_1 < r_2$, let $E = \{ru :\ r_1 < r < r_2,\ u \in A\}$. Show the equality for $1_E$, and then pass to characteristics of Borel sets.

(2.7.18) Show that, when $n = 2$, the equality (2.43) becomes the usual Calculus polar coordinate change of variable; that is, for $f \in C_c(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2} f\, dm = \int_0^\infty \int_0^{2\pi} r\, f(r\cos t, r\sin t)\, dt\, dr.$$

(2.7.19) *(Volume of a ball in $\mathbb{R}^n$)* Let $B_r(0)$ be the ball of radius $r$, in $\mathbb{R}^n$, centered at the origin.

(i) Show that $m(B_1(0)) = \sigma(S_{n-1})/n$.

(ii) Calculate $\int_{\mathbb{R}^n} e^{-|x|^2}\, dm$ using Fubini.

(iii) Calculate $\int_{\mathbb{R}^n} e^{-|x|^2}\, dm$, using (2.43), in terms if $\sigma(S_{n-1})$ and the Gamma Function from Exercise 3.1.6.

(iv) Show that $m(B_r(0)) = r^n\, m(B_1(0))$.

(v) Use what you found to write a formula for $m(B_r(0))$.

(vi) Let $C_n \subset \mathbb{R}^n$ denote the hypercube of side 2 centered at the origin, and $B_n \subset \mathbb{R}^n$ the ball of radius 1 centered at the origin. Show that $\lim_{n\to\infty} \frac{m(B_n)}{m(C_n)} = 0$. This result is not so anti-intuitive when you notice that $\operatorname{diam} B_n = 2$, while $\operatorname{diam} C_n = 2\sqrt{n}$. Another way of seeing it as natural is that the unit ball cannot touch the areas near the corners of the cube. The square has four corners, the cube has eight corners, and in dimension $n$ the hypercube has $2^n$ corners. So as dimension grows, there are more an more parts of the cube that cannot be touched by the ball.

(2.7.20) Let $\alpha > 0$, $n \in \mathbb{N}$. In $\mathbb{R}^n$, evaluate

$$\int_{|x|\leq 1} \frac{1}{|x|^\alpha}\, dx, \qquad \text{and} \qquad \int_{|x|\geq 1} \frac{1}{|x|^\alpha}\, dx.$$

(2.7.21) Let $(X, \mathcal{A}, \mu)$ be a $\sigma$-finite measure space and $f : X \to \mathbb{C}$ measurable. Let $\omega_f : [0, \infty) \to [0, \infty)$ be the distribution function

$$\omega_f(t) = \mu(\{|f| > t\}).$$

Prove that

$$\int_X |f|^p = \int_0^\infty p\, t^{p-1}\, \omega_f(t)\, dt, \qquad 1 \le p < \infty.$$

More generally, show that if $\gamma : [0, \infty) \to [0, \infty)$ is increasing, differentiable, and $\gamma(0) = 0$, then

$$\int_X \gamma \circ f \, d\mu = \int_0^\infty \omega_f(t)\, \gamma'(t)\, dt.$$

## 2.8. $L^p$-Spaces

The $L^p$ spaces appear naturally in many areas of mathematics, from differential equations to Banach Spaces. Before we begin to define things, there is a simple inequality that we will need almost immediately.

***2.8.1. Lemma. (An easy inequality)***

*Let $a, b \in [0, \infty)$, $p \in [0, \infty)$. Then*

$$(a + b)^p \le 2^p\, (a^p + b^p). \tag{2.44}$$

*Proof.* Assume without loss of generality that $a \ge b$. Then

$$(a + b)^p = a^p \left(1 + \frac{b}{a}\right)^p \le a^p\, 2^p \le a^b\, 2^p \left(1 + \frac{b^p}{a^p}\right) = 2^p\, (a^p + b^p). \qquad \square$$

**2.8.2. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space and $p \in [0, \infty]$. Let*

$$\mathscr{L}^p(X, \mathcal{A}, \mu) = \left\{ f : X \to \mathbb{C} : \quad \textit{measurable}, \int_X |f|^p\, d\mu < \infty \right\}, \quad p < \infty,$$

$$\mathscr{L}^\infty(X, \mathcal{A}, \mu) = \{ f : X \to \mathbb{C} : \quad \textit{measurable}, \exists c > 0 : \; |f| < c \;\; \text{a.e.} \}.$$

The case $p = 0$ is fully non-interesting, as we get either all measurable functions if $\mu(X) < \infty$, or just the zero function if $\mu(X) = \infty$. We note that $\mathscr{L}^1(X, \mathcal{A}, \mu)$ is the set of integrable functions for the measure space $(X, \mathcal{A}, \mu)$.

All these are vector spaces: it is apparent that $\mathscr{L}^\infty(X, \mathcal{A}, \mu)$ is a vector space—just use the triangle inequality. For $p < \infty$, the inequality (2.44) gives us that $|f + \lambda g|^p \leq 2^p \left(|f|^p + |\lambda|^p |g|^p\right)$ and it follows $\mathscr{L}^p(X, \mathcal{A}, \mu)$ is a vector space.

Associated with the $\mathscr{L}^p$-spaces are the $p$-**norms**.

### 2.8.3. Definition.

*If $p \in [1, \infty)$ and $f \in \mathscr{L}^p(X)$, then the $p$-norm of $f$ is the number*

$$\|f\|_p = \left( \int_X |f|^p \, d\mu \right)^{1/p}.$$

*If $f \in \mathscr{L}^\infty(X)$, the **infinity norm** of $f$ is the number*

$$\|f\|_\infty = \operatorname{ess\,sup} |f|,$$

*where the **essential supremum** of $f$ is*

$$\operatorname{ess\,sup} |f| = \inf\{c > 0 : \ |f| < c \ \ \text{a.e.}\}.$$

The notion of norm appears in Chapters 4 and 5 and pretty much everywhere else in this book, with the exception of Chapter 3. For the reader who has not yet gone through the aforementioned chapters nor has some background on normed spaces, let us recall:

### 2.8.4. Definition.

*A **norm** on a real or complex vector space $V$ is a map $n : V \to [0, \infty)$ such that*

*(i) $n(v) = 0$ if and only if $v = 0$;*

*(ii) $n(\lambda v) = |\lambda| \, n(v)$ for all scalar $\lambda$ (**homogeneity**);*

*(iii) $n(v + w) \leq n(v) + n(w)$ for all $v, w \in V$ (**triangle inequality**).*

The primary interest of having a norm on a vector space is that $d(v, w) = n(v - w)$ defines a translation-invariant metric on $V$, and the vector operations are continuous in this metric. The canonical example of this is $V = \mathbb{R}^n$, where $n$ is the Euclidean norm

$$n(x) = \|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}. \tag{2.45}$$

The norm of an element $v$ is most commonly denoted by $\|v\|$, possibly with a subindex—as in (2.45)—to distinguish it from other norms.

It should be apparent that the $p$-norms defined above are not norms. The problem is that neither the condition $\int_X |f|^p \, d\mu = 0$ nor $\operatorname{ess\,sup} |f| = 0$ guarantee that $f = 0$; just that $f = 0$ a.e. As functions that agree almost everywhere are indistinguishable via integrals, we consider an equivalence relation

$$f \sim g \qquad \text{if and only if} \qquad f = g \ \text{ a.e.} \tag{2.46}$$

This is an equivalence relation that is compatible with the pointwise sum and product of functions, and with the $p$-norms (Exercise 2.8.1). We then define

**2.8.5. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $\sim$ the relation from (2.46), and $p \in [1, \infty]$. Then*

$$L^p(X, \mathcal{A}, \mu) = \mathscr{L}^p(X, \mathcal{A}, \mu)/\sim .$$

The quotient in Definition 2.8.5 can also be seen as the quotient by the subspace $N = \{f : X \to \mathbb{C} : f = 0 \ \text{ a.e.}\}$. In $L^p(X, \mathcal{A}, \mu)$ we now have that $\|f\|_p = 0$ implies $f = 0$ (as classes). One can also show that $|f| \leq \|f\|_\infty$ a.e. (Exercise 2.8.3).

Because the usual operations with classes are well-defined, in practice we forget that the elements of $L^p$ are classes and we treat them as functions. The only time one needs to be careful is when evaluating functions pointwise, as this is **not** well-defined: given any $f \in L^p(X, \mathcal{A}, \mu)$, as long as $x$ is not an atom we can redefine $f$ at $x$ to take any value and the new function will still be in the same class. For instance, the map $f \longmapsto f(0)$ makes no sense in $L^p(\mathbb{R}, \mathcal{M}(\mathbb{R}), m)$; it is not well-defined.

Another case where paying attention to representatives makes sense is the following. We do allow $f \in L^p(X)$ to take the values $\pm\infty$ when it is real-valued. But the fact that $\|f\|_p < \infty$ implies that $\{|f| = \infty\}$ has to be a nullset (Exercise 2.8.2). So we can always change $f$ to be zero, say, on this nullset and that way we obtain a representative of the class of $f$ that is finite everywhere. This means that in practice we can always assume that $f \in L^p(X)$ is finite everywhere.

When $\mathcal{A} = \mathcal{P}(X)$ and $\mu$ is the counting measure, we get the space $\ell^p(X)$, the most common case being $\ell^p(\mathbb{N})$. In other words,

**2.8.6. Definition.**

$$\ell^p(\mathbb{N}) = \Big\{a : \mathbb{N} \to \mathbb{C} : \ \sum_{n=1}^\infty |a_n|^p < \infty\Big\},$$

*and*

$$\ell^\infty(\mathbb{N}) = \{a : \mathbb{N} \to \mathbb{C} : \ \sup_n |a_n| < \infty\}.$$

In these cases the measure space has no non-trivial nullsets, so no quotient is needed. A very familiar case arises when one considers $X = \{1, \dots, n\}$, $\mathcal{A} = \mathcal{P}(X)$ and $\mu$ the counting measure. In this situation $L^2(X, \mathcal{A}, \mu)$, if we consider real-valued functions, is precisely $\mathbb{R}^n$ with the Euclidean norm (and $\mathbb{C}^n$ with the Euclidean norm if one considers complex-valued functions). In some rare cases we will consider $\ell^p(X)$ for a set $X$ other than $\mathbb{N}$.

Most of the time we consider $p \geq 1$; these spaces are the "good ones"; the cases where $p < 1$ are rather pathological—they will be briefly discussed later in subsection 5.6.7. Let us note for now that the "$p$-norm" is not a norm when $p < 1$. For $p \geq 1$, we will prove the triangle inequality in Corollary 2.8.10.

We will soon see that the natural metric induced by the $p$-norm makes the space $L^p(X, \mathcal{A}, \mu)$ a complete metric space.

It is common to write, depending on the circumstances, both $L^p(X)$ and $L^p(\mu)$ instead of $L^p(X, \mathcal{A}, \mu)$. For instance, for $X \subset \mathbb{R}$ with Lebesgue measure, $L^p(X)$ is standard; if one sees $L^p(\mathbb{R}^2)$ or $L^p[0,1]$ without further clarification, the assumption is that the Lebesgue measure is used.

A common feature in the world of $L^p$-spaces is the **conjugate exponent**: given $p \in [1, \infty]$ its conjugate exponent is $q$ with $\frac{1}{p} + \frac{1}{q} = 1$. That is, $q = \infty$ when $p = 1$ and otherwise $q = \frac{p}{p-1}$. When $p$ is mentioned and $q$ appears without further clarification, it is assumed that $q$ is the conjugate exponent of $p$. The most common cases one uses are $p = 1, 2, \infty$ (where $q$ is, respectively, $\infty, 2, 1$), but all other cases, like $p = 5.7$ and $q = 5.7/4.7$ or $p = \sqrt{2}$ and $q = \frac{\sqrt{2}}{\sqrt{2}-1}$, are allowed unless explicitly restricted.

***2.8.7. Lemma. (Young's Inequality for positive real numbers)***

*Let $x, y \in \mathbb{R}$, $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then $|xy| \leq \dfrac{|x|^p}{p} + \dfrac{|y|^q}{q}$.*

*Proof.* We note that it is enough to prove the inequality for $x > 0$, $y > 0$, since $xy = 0$ automatically makes the inequality true, and $|xy| = |x|\,|y|$. Since log is concave downwards and $1/p + 1/q = 1$, we have

$$\begin{aligned}\log(xy) = \log((x^p)^{1/p}(y^q)^{1/q}) &= \frac{1}{p}\log(x^p) + \frac{1}{q}\log(y^q) \\ &\leq \log\left(\frac{x^p}{p} + \frac{y^q}{q}\right).\end{aligned}$$

Exponentiating, we get the result. $\square$

The Hölder and Minkowski inequalities are the basic tools of the trade here.

**2.8.8. Theorem. (Hölder Inequality)**

*Let* $(X, \mathcal{A}, \mu)$ *a measure space,* $1 \le p \le \infty$ *and* $q$ *its conjugate exponent,* $f \in L^p(X)$*,* $g \in L^q(X)$*. Then*

$$\|fg\|_1 \le \|f\|_p \, \|g\|_q.$$

*Equality occurs precisely when either* $f = 0$*,* $g = 0$*, or when there exists* $\alpha \in (0, \infty)$ *with* $|f|^p = \alpha |g|^q$*.*

*Proof.* If $\|f\|_p = 0$ or $\|g\|_q = 0$, then one of them is zero almost everywhere, and hence both sides of the inequality are zero. So we only need to consider the case where $\|f\|_p \|g\|_q > 0$.

Assume first that $1 < p < \infty$. Assume further that $\|f\|_p = \|g\|_q = 1$. Then, applying Young's inequality pointwise, we have

$$\|fg\|_1 = \int_X |fg| \, d\mu \le \int_X \left( \frac{|f|^p}{p} + \frac{|g|^q}{q} \right) d\mu = \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Now, for general $f, g$ we apply the previous inequality to $f / \|f\|_p$, $g / \|g\|_q$ and we obtain the result.

When $p = 1$, $q = \infty$, we have

$$\int_X |fg| \, d\mu = \int_X |f| \, |g| \le \int_X |f| \, \|g\|_\infty \, d\mu = \|f\|_1 \, \|g\|_\infty. \qquad \square$$

Exchanging the roles of $f$ and $g$ gives us the case $p = \infty$, $q = 1$. For the case of equality we refer to Exercise 2.8.11.

The case $p = q = 2$ of Hölder's Inequality is most often called the **Cauchy–Schwarz Inequality**.

**2.8.9. Corollary. (Discrete Hölder Inequality)**

*Let* $1 \le p \le \infty$*,* $f \in \ell^p(\mathbb{N})$*,* $g \in \ell^q(\mathbb{N})$*. Then*

$$\sum_k |f(k) g(k)| \le \|f\|_p \, \|g\|_q.$$

*Proof.* Apply Theorem 2.8.8 to the measure space $\mathbb{N}$ with the counting measure. $\square$

The Minkowski inequality is nothing but the triangle inequality for the $p$-norm.

***2.8.10. Corollary. (Minkowski Inequality)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $1 \le p \le \infty$, and $f, g \in L^p(X)$. Then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{2.47}$$

*Equality in (2.47) in the case $p < \infty$ occurs precisely when there exist $\alpha, \beta \ge 0$ with $\alpha f = \beta g$ a.e.*

*Proof.* We do $p = \infty$ first. Suppose that $|f| \le c$ and $|g| \le d$ a.e. Then $|f+g| \le |f| + |g| \le c + d$ a.e. This shows that $\|f+g\|_\infty \le c+d$; as we can choose $c$ and $d$ arbitrarily close to $\|f\|_\infty$ and $\|g\|_\infty$ respectively, we get that $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

Now assume $p < \infty$. When $\|f+g\|_p = 0$, the inequality holds trivially. So we assume that $\|f+g\|_p > 0$. Using first the triangle inequality, then Hölder, and then that $q(p-1) = p$ for $p > 1$ (note that when $p = 1$ we will not need this number),

$$\begin{aligned}
\|f+g\|_p^p &= \int_X |f+g|^p \, d\mu = \int_X |f+g| \, |f+g|^{p-1} \, d\mu \\
&\le \int_X |f| \, |f+g|^{p-1} + \int_X |g| \, |f+g|^{p-1} \\
&\le \|f\|_p \left( \int_X |f+g|^{q(p-1)} \, d\mu \right)^{1/q} + \|g\|_p \left( \int_X |f+g|^{q(p-1)} \, d\mu \right)^{1/q} \\
&= (\|f\|_p + \|g\|_p) \, \|f+g\|_p^{p/q}
\end{aligned}$$

Now the fact that $p - \frac{p}{q} = 1$, together with $\|f+g\|_p > 0$, gives us (2.47). For the case of equality we refer to Exercise 2.8.12. $\square$

At first sight it is not clear why the infinity norm is bundled with the $p$-norms. But the name is deserved:

***2.8.11. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{C}$ measurable. If $\|f\|_r < \infty$ for some $r > 1$, then*

$$\|f\|_\infty = \lim_{p \to \infty} \|f\|_p.$$

*Proof.* If $\|f\|_\infty = 0$, then $f = 0$ a.e. and the equality is trivial. If $0 < \|f\|_\infty < \infty$, fix $\delta > 0$ with $\delta < \|f\|_\infty$. Let $S_\delta = \{|f| \ge \|f\|_\infty - \delta\}$. We have $0 < \mu(S_\delta)$

by definition of $\|f\|_\infty$, and $\mu(S_\delta) < \infty$, since otherwise $\|f\|_r = \infty$. Then

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p} \geq \left(\int_{S_\delta} |f|^p \, d\mu\right)^{1/p} \geq \big(\|f\|_\infty - \delta\big)\, \mu(S_\delta)^{1/p}.$$

Since $\mu(S_\delta) > 0$, as $p$ goes to infinity we get $\liminf_{p\to\infty} \|f\|_p \geq \|f\|_\infty - \delta$. As this can be done for all $\delta$,

$$\liminf_{p\to\infty} \|f\|_p \geq \|f\|_\infty.$$

On the other hand, for $p > r$ we have

$$\|f\|_p = \left(\int_X |f|^{p-r}\, |f|^r \, d\mu\right)^{1/p} \leq \|f\|_\infty^{(p-r)/p}\, \|f\|_r^{r/p} = \|f\|_\infty^{1-r/p}\, \|f\|_r^{r/p}.$$

As $p \to \infty$, we obtain

$$\limsup_{p\to\infty} \|f\|_p \leq \|f\|_\infty.$$

So $\liminf_{p\to\infty} \|f\|_p = \limsup_{p\to\infty} \|f\|_p = \|f\|_\infty$, showing that the limit exists and is $\|f\|_\infty$.

Finally, consider the case where $\|f\|_\infty = \infty$. Fix $c > 0$. The fact that the infinity norm of $f$ is infinite guarantees that $|f| \geq c$ on a set $S$ with $\mu(S) > 0$. Then

$$\|f\|_p = \left(\int_X |f|^p\right)^{1/p} \geq \left(\int_S |f|^p\right)^{1/p} \geq c\, \mu(S)^{1/p}.$$

As this can be done for any $p$,

$$\liminf_{p\to\infty} \|f\|_p \geq c.$$

And $c > 0$ can be chosen arbitrarily, so we get that $\liminf_{p\to\infty} \|f\|_p = \infty$. This shows that the limit exists and is infinity, as we wanted to show. □

Recall that a metric space is **complete** if every Cauchy sequence is convergent. It is common to attribute the result below to Riesz–Fischer, although for some the Riesz–Fischer theorem is about convergence of Fourier series.

***2.8.12. Theorem. (Riesz–Fischer)***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $p \in [1, \infty]$. Then $L^p(X, \mathcal{A}, \mu)$ is complete.*

*Proof.* Let $\{f_n\} \subset L^p$ be a Cauchy sequence. This means that for all $\varepsilon > 0$ there exists $n_0$ such that $\|f_n - f_m\|_p < \varepsilon$ whenever $n, m \geq n_0$. We want to show that $\{f_n\}$ is convergent.

We consider first the case $p = \infty$. For each $m, n \in \mathbb{N}$ there exists $A_{m,n} \in \mathcal{A}$, with $\mu(A_{m,n}) = 0$, and such that $|f_n - f_m| \leq \|f_n - f_m\|_\infty$ outside of $A_{m,n}$. There is also $A_n \in \mathcal{A}$, with $\mu(A_n) = 0$ and such that $|f_n| \leq \|f_n\|_\infty$ outside of $A_n$. Let

$$A = \bigcup_n A_n \cup \bigcup_{m,n} A_{m,n}.$$

Being a countable union of nullsets, $A$ is itself a nullset. For all $x \in X \setminus A$, the sequence $\{f_n(x)\}$ is uniformly Cauchy; that is, given $\varepsilon > 0$ there exists $n_0$ such that $|f_n(x) - f_m(x)| < \varepsilon$ whenever $m, n \geq n_0$. By the completeness of $\mathbb{C}$, the limit $f(x) = \lim_n f_n(x)$ exists for all $x \in X \setminus A$. Define $f(x) = 0$ for $x \in A$. Given $x \in X \setminus A$ and $\varepsilon > 0$, choose $n_0$ as above. For $m, n \geq n_0$,

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| \leq |f(x) - f_m(x)| + \varepsilon.$$

Taking the limit as $m \to \infty$, we get that $|f(x) - f_n(x)| \leq \varepsilon$ for all $x \in X \setminus A$. Thus, $\|f - f_n\|_\infty < \varepsilon$, proving that $f_n \to f$ in $L^\infty$. This also implies that $f \in L^\infty$, for

$$|f(x)| \leq |f(x) - f_n(x)| + \|f_n\|_\infty.$$

By the Reverse Triangle Inequality (Exercise 1.8.24), $\{\|f_n\|_\infty\}$ is Cauchy, hence bounded. Thus $|f(x)| \leq |f(x) - f_n(x)| + c$ a.e. for all $n$, and taking limit we get $|f| \leq c$ a.e., showing that $\|f\|_\infty < \infty$.

When $1 \leq p < \infty$, choose $n_1$ such that $\|f_n - f_{n_1}\|_p < 1/2$ for all $n \geq n_1$. Inductively, choose $n_k \geq n_{k-1}$ such that $\|f_n - f_{n_k}\|_p < 2^{-k}$ for all $n \geq n_k$. We want to define

$$f = f_{n_1} + \sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k})$$

(note that this series is telescopic, so $f$ is necessarily the pointwise limit of $f_{n_k}$, if it exists). For this, we need to show that the series converges pointwise almost everywhere. We have

$$\Big\| \sum_{k=1}^{N} |f_{n_{k+1}} - f_k| \Big\|_p \leq \sum_{k=1}^{N} \|f_{n_{k+1}} - f_{n_k}\|_p \leq \sum_{k=1}^{N} 2^{-k} \leq 1.$$

Thus, by Monotone Convergence,

$$\begin{aligned}
\int_X \Big( \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| \Big)^p d\mu &= \lim_{N \to \infty} \int_X \Big( \sum_{k=1}^{N} |f_{n_{k+1}} - f_{n_k}| \Big)^p d\mu \\
&= \lim_{N \to \infty} \Big\| \sum_{k=1}^{N} |f_{n_{k+1}} - f_k| \Big\|_p^p \leq 1.
\end{aligned}$$

Then $\displaystyle\sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| < \infty$ a.e. and $f$ is well-defined. Moreover,

$$\|f\|_p \leq \|f_{n_1}\|_p + \Big\| \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| \Big\|_p \leq \|f_{n_1}\|_p + 1,$$

so $f \in L^p(X)$. An interesting property of the definition of $f$ being telescopic is that

$$f = f_{n_h} + \sum_{k=h}^{\infty} (f_{n_{k+1}} - f_{n_k})$$

for any $h \in \mathbb{N}$. Then

$$\|f - f_{n_h}\|_p = \left\| \sum_{k=h}^{\infty} f_{n_{k+1}} - f_{n_k} \right\|_p \leq \sum_{k=h}^{\infty} 2^{-k} = 2^{-h+1},$$

showing that $f_{n_k} \to f$ in $L^p(X)$. As $\{f_{n_k}\}$ is a convergent subsequence of a Cauchy sequence, the full sequence converges to the limit $f$. □

We state explicitly a fact that was obtained in the proof of Theorem 2.8.12:

***2.8.13. Corollary.***

*Let $X$ be a measure space, $p \in [1, \infty]$, and $\{f_n\} \subset L^p(X)$ a sequence that converges to $f \in L^p(X)$. Then there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \to f$ pointwise a.e.*

*Proof.* For $p < \infty$ the subsequence was constructed in the proof of Theorem 2.8.12. For $p = \infty$, we showed that the original sequence is legit pointwise Cauchy out of a nullset, so it converges pointwise a.e. □

Next we offer several density results. They are used often, since one can do little with an arbitrary element in $L^p(X)$. The results below do work for $p \in (0, 1)$; but $L^p$ for $p < 1$ is a different beast—this will be discussed starting on subsection 5.6.7—and we prefer to emphasize the well-behaved spaces for now.

***2.8.14. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $p \in [1, \infty)$. Then the functions in $L^p(X)$ whose support has finite measure are dense in $L^p(X)$.*

*Proof.* Fix $f \in L^p(X, \mathcal{A}, \mu)$. Let $X_n = \left\{|f| > \frac{1}{n}\right\}$. We have

$$\mu(X_n) = \int_{X_n} 1 \, d\mu = n^p \int_{X_n} \frac{1}{n^p} \, d\mu \leq n^p \int_{X_n} |f|^p \, d\mu \leq n^p \|f\|_p^p < \infty.$$

The sequence $\{X_n\}_n$ is increasing, and $\{|f| > 0\} = \bigcup_n X_n$. So, using Monotone Convergence (Exercise 2.5.19),

$$\|f\|_p^p = \int_X |f|^p \, d\mu = \int_{\{|f|>0\}} |f|^p \, d\mu = \lim_n \int_{X_n} |f|^p \, d\mu.$$

Then

$$\begin{aligned}\|f - f\,1_{X_n}\|_p^p &= \int_X |f(1 - 1_{X_n})|^p \, d\mu = \int_X |f|^p \,(1 - 1_{X_n})\, d\mu \\ &= \|f\|_p^p - \int_{X_n} |f|^p \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

□

*2.8.15. Remark.* The exclusion of the case $p = \infty$ in Proposition 2.8.14 was not arbitrary. Consider $1 \in \ell^\infty(\mathbb{N})$. If $g \in \ell^\infty(\mathbb{N})$ is supported on a set of finite measure, since the measure is the counting measure this means that $g$ is supported on a finite set. So there exists $m$ with $g(m) = 0$. Hence $\|1 - g\|_\infty \geq |1 - g(m)| = 1$, showing that 1 cannot be approximated with functions of finite support.

***2.8.16. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $p \in [1, \infty)$. Then the set of bounded functions with finite-measure support is dense in $L^p(X)$. In particular, the bounded functions are dense in $L^p(X)$.*

*Proof.* Fix $f \in L^p(X, \mathcal{A}, \mu)$ and let $\varepsilon > 0$. By Proposition 2.8.14 there exists $f_1 \in L^p(X, \mathcal{A}, \mu)$, finitely supported, with $\|f - f_1\|_p < \varepsilon/2$. Let $E = \{|f_1| > 0\}$ and $E_n = \{0 < |f_1| < n\}$. We have $\mu(E) < \infty$ and $E = \bigcup_n E_n$. As the union is non-decreasing, $\lim_n \mu(E_n) = \mu(E)$, and $1_{E_n} \nearrow 1_E$ pointwise. By Monotone Convergence (Exercise 2.5.19),

$$\lim_n \int_{E_n} |f_1|^p \, d\mu = \int_E |f_1|^p \, d\mu,$$

and then

$$\|f_1 - f_1\,1_{E_n}\|_p^p = \int_{E\setminus E_n} |f_1|^p = \int_E |f_1|^p \, d\mu - \int_{E_n} |f_1|^p \, d\mu \xrightarrow[n\to\infty]{} 0.$$

So we may choose $n$ such that $\|f_1 - f_1\,1_{E_n}\|_p < \varepsilon/2$ and then $\|f - f_1\,1_{E_n}\|_p < \varepsilon$. The function $f_1\,1_{E_n}$ is bounded and supported on a set of finite measure.

□

Proposition 2.8.16 also fails when $p = \infty$ and $\mu(X) = \infty$. For if $\mu(X) = \infty$ and $f \in L^\infty(X)$, then $\|f - 1\|_\infty \geq 1$ if $\mu(\{|f| > 0\}) < \infty$. This is because $|f - 1| = 1$ on $X \setminus \{|f| > 0\}$, which is a set of infinite (hence positive) measure. When $\mu(X) < \infty$ both propositions hold trivially in $L^\infty(X)$ in the sense that every $f \in L^\infty$ is almost everywhere a bounded function with finite-measure support.

***2.8.17. Proposition.***

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $p \in [1, \infty]$. Then the simple functions are dense in $L^p(X)$.*

*Proof.* Let $f \in L^p(X, \mathcal{A}, \mu)$ and fix $\varepsilon > 0$. When $p = \infty$, we get directly from Theorem 2.4.13 that there exists $s_1$ simple with $\|f - s_1\|_\infty < \varepsilon$.

When $p < \infty$, by Proposition 2.8.16 there exists $X_0$ with $\mu(X_0) < \infty$ and bounded $f_0 \in L^p$ with $f_0 = f_0 1_{X_0}$ and $\|f - f_0\|_p < \varepsilon/2$. By Theorem 2.4.13 there exists $s_1$ simple with $\|f_0 - s_1\|_\infty < \varepsilon/\big(2\mu(X_0)^{1/p}\big)$, and $s_1 = 0$ outside of $X_0$. Then

$$\int_X |f_0 - s_1|^p = \int_{X_0} |f_0 - s_1|^p \leq \frac{\varepsilon^p}{2^p \mu(X_0)} \int_{X_0} 1 = \frac{\varepsilon^p}{2^p},$$

and

$$\|f - s_1\|_p \leq \|f - f_0\|_p + \|f_0 - s_1\|_p \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

***2.8.18. Proposition.***

*Let $X$ be a locally compact Hausdorff space and $\mu$ Radon measure on $X$. Let $p \in [1, \infty)$. Then $C_c(X)$ is dense in $L^p(X, \mathcal{B}(X), \mu)$.*

*Proof.* Let $f \in L^p(X, \mathcal{B}(X), \mu)$. By Proposition 2.8.16 we may assume that $\mu\{|f| > 0\} < \infty$ and that $|f| \leq c$ for some $c$. Fix $\varepsilon > 0$. By Lusin's Theorem (Corollary 2.6.13), applied on the finite-measure space $X \cap \{|f| > 0\}$ there exists $g \in C_c(X)$ with $|g| \leq c$ and such that $\mu(\{f \neq g\}) < \varepsilon^p/(2c)^p$. Then

$$\|f - g\|_p = \left( \int_{\{f \neq g\}} |f - g|^p \right)^{1/p} \leq 2c\mu(\{f \neq g\})^{1/p} < \varepsilon. \qquad \square$$

Since a uniform limit of continuous functions is continuous, we cannot expect to get a version of Proposition 2.8.18 for $p = \infty$.

When the space has more structure, more can be proved. Note that in general a "locally compact space" is just a topological space, there is no vector space structure. But below we also need it to be a vector space for the translation to be defined. When we say that $X$ is a *locally compact topological vector space*, we mean that $X$ is a vector space, that it has a locally compact Hausdorff topology, and that the vector space operations are continuous. We refer to Section 5.4 for details on topological vector spaces; for now, if this feels too abstract, it is enough to assume that $X = \mathbb{R}^d$ and that the open sets in the proof are balls.

**2.8.19. Lemma. ($L^p$-continuity)**

*Let $X$ be a locally compact Hausdorff topological vector space and $\mu$ an outer and inner regular Borel measure on $X$. Let $p \in [1, \infty)$ and $f \in L^p(X, \mu)$. For $t \in X$, let $f_t(x) = f(x - t)$. Then*

$$\lim_{t \to 0} \|f_t - f\|_p = 0. \tag{2.48}$$

*Proof.* Exercise 2.8.24. □

To work with convolutions we will need the integral version of Minkowski's Inequality. Thinking of integrals as a generalization of sums, the inequality is a generalization of the triangle inequality for the $p$-norm ("the norm of the sum is less than or equal the sum of the norms"), which we proved in Corollary 2.8.10.

**2.8.20. Proposition. (Minkowski's Integral Inequality)**

*Let $(X, \mu)$ and $(Y, \nu)$ be $\sigma$-finite measure spaces, and $f : X \times Y \to \mathbb{C}$ be measurable, $p \in [1, \infty)$. Then*

$$\left( \int_Y \left| \int_X f(x, y)\, d\mu(x) \right|^p d\nu(y) \right)^{1/p} \leq \int_X \left( \int_Y |f(x, y)|^p d\nu(y) \right)^{1/p} d\mu(x). \tag{2.49}$$

*Proof.* We consider first the case where $\mu(X) < \infty$ and $\nu(Y) < \infty$ and then extend at the end. Since

$$\left( \int_Y \left| \int_X f(x, y)\, d\mu(x) \right|^p d\nu(y) \right)^{1/p} \leq \left( \int_Y \left[ \int_X |f(x, y)|\, d\mu(x) \right]^p d\nu(y) \right)^{1/p}$$

and we will prove that the right-hand-side in (2.49) is larger, we may assume without loss of generality that $f \geq 0$. Assume furthermore that $f$ is bounded.

Let us denote the left-hand-side in (2.49) by $C$ (without absolute value, for we are assuming that $f \geq 0$). If $C = 0$, we have nothing to prove, as (2.49) holds trivially. We also have $C < \infty$ by our current assumptions, since $0 \leq f \leq k$ for some $k > 0$, and $\mu(X) < \infty$ and $\nu(Y) < \infty$. The argument below is essentially the same as in Corollary 2.8.10. Using Tonelli and Hölder,

$$\begin{aligned}
C^p &= \int_Y \left[\int_X f(x,y)\,d\mu(x)\right]^{p-1} \left(\int_X f(x,y)\,d\mu(x)\right) d\nu(y) \\
&= \int_Y \int_X \left[\int_X |f(x,y)|\,d\mu(x)\right]^{p-1} f(x,y)\,d\mu(x)\,d\nu(y) \\
&= \int_X \int_Y \left[\int_X |f(x,y)|\,d\mu(x)\right]^{p-1} f(x,y)\,d\nu(y)\,d\mu(x) \\
&\le \int_X \left(\int_Y \left[\int_X f(x,y)\,d\mu(x)\right]^{q(p-1)} d\nu(y)\right)^{1/q} \left(\int_Y \left[f(x,y)\right]^p d\nu(y)\right)^{1/p} d\mu(x) \\
&= \int_X \left(\int_Y \left[\int_X f(x,y)\,d\mu(x)\right]^{p} d\nu(y)\right)^{1/q} \left(\int_Y \left[f(x,y)\right]^p d\nu(y)\right)^{1/p} d\mu(x).
\end{aligned}$$

The first bracket in the last integral above is $C^{p/q}$; dividing by it on both sides, and noting that $p - \frac{p}{q} = 1$, we obtain

$$\left(\int_Y \left[\int_X f(x,y)\,d\mu(x)\right]^{p} d\nu(y)\right)^{1/p} \le \int_X \left(\int_Y \left[f(x,y)\right]^p d\nu(y)\right)^{1/p} d\mu(x)$$

as desired. For the general case, we may write $X = \bigcup_n X_n$, $Y = \bigcup_n Y_n$ increasing unions of finite-measure sets. We can also write $f = \lim f_n$, a monotone limit of bounded functions (namely, take $f_n = f\,1_{\{|f|\le n\}}$). Then the above argument applies to the functions $f_n\,1_{X_n\times Y_n}$, and we get (2.49) by Monotone Convergence. $\square$

The example in Remark 2.7.18, with the roles of $\mu$ and $\nu$ exchanged, shows that $\sigma$-finiteness is crucial in Proposition 2.8.20 (for any $p \in [1,\infty)$).

**2.8.21. Definition.**

*Consider $X = \mathbb{R}^d$ with Lebesgue measure. For good enough functions $f, g$ so that the integral makes sense, the* ***convolution*** *of $f$ and $g$ is the function*

$$(f * g)(x) = \int_{\mathbb{R}^d} f(t)\,g(x-t)\,dt.$$

The spaces we require here for each of the functions are chosen according to our needs, but of course one can define the convolution as long as the integral above makes sense. Via the change of variable $u = x - t$, we see that $f * g = g * f$ (Exercise 2.7.13). Though we will not discuss it here, any locally compact topological group admits a unique (up to a positive multiple) so-called **Haar measure** on its Borel $\sigma$-algebra, which is a left-translation invariant Radon measure;

among many other things this allows one to define convolution integrating over the group. The Lebesgue measure is the normalized Haar measure for $\mathbb{R}$. For a discrete group, the Haar measure is the counting measure.

We mention that the inequality in Proposition 2.8.22 is a particular easy case of the well-known **Young's Convolution Inequality** (Exercise 2.8.17):

$$\text{if} \qquad \frac{1}{p}+\frac{1}{q}=\frac{1}{r}+1, \qquad \text{then} \qquad \|f*g\|_r \le \|f\|_p\,\|g\|_q. \tag{2.50}$$

We will use the fairly standard notation $D_\alpha$ for the partial derivative operator corresponding to the multi-index $\alpha$. And $C_c^k(\mathbb{R}^d)$ denotes the space of $k$-times continuously differentiable functions with compact support on $\mathbb{R}^d$.

***2.8.22. Proposition.***

*Let $f \in L^p(\mathbb{R}^d)$, $g \in L^1(\mathbb{R}^d)$. Then*

*(i)* $\|f*g\|_p \le \|f\|_p\|g\|_1$; *in particular,* $f*g \in L^p(\mathbb{R}^d)$;

*(ii) if* $g \in C_c^k(\mathbb{R}^d)$, *then* $D_\alpha(f*g) = f*D_\alpha g$ *for all* $|\alpha| \le k$ *and so* $f*g \in C^k(\mathbb{R}^d)$;

*(iii) if* $\displaystyle\int_{\mathbb{R}^d}|g| = 1$ *and* $g_n = n^d g(nx)$, *then* $\lim_n \|f - f*g_n\|_p = 0$.

*Proof.*

(i) We have, using Minkowski's Integral Inequality,

$$\begin{aligned}
\|f*g\|_p &= \left(\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d} f(x-t)\,g(t)\,dt\right|^p dx\right)^{1/p} \\
&\le \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|f(x-t)|^p\,|g(t)|^p\,dx\right)^{1/p} dt \\
&= \int_{\mathbb{R}^d}|g(t)|\left(\int_{\mathbb{R}^d}|f(x-t)|^p\ dx\right)^{1/p} dt \\
&= \|f\|_p\,\|g\|_1.
\end{aligned}$$

(ii) If $e_1,\dots,e_d$ denotes the canonical basis in $\mathbb{R}^d$, the Newton quotient for the $j^{\text{th}}$ partial derivative of $f*g$ at a fixed $x$ is

$$\begin{aligned}
N_j(x,h) &= \frac{(f*g)(x+he_j)-(f*g)(x)}{h} \\
&= \int_{\mathbb{R}^d} f(t)\,\frac{g(x+he_j-t)-g(x-t)}{h}\,dt \\
&= \int_{\mathbb{R}^d} f(t)\,D_j g(x-t+\gamma_{t,h}e_j)\,dt
\end{aligned}$$

by using the Mean Value Theorem on the function $s \longmapsto g(x - t + se_j)$. The numbers $\gamma_{t,h}$ satisfy $|\gamma_{t,h}| \le h$. Now, because $g$ is continuously differentiable with compact support, $D_j g$ is continuous with compact support and hence uniformly continuous. So for fixed $\varepsilon > 0$ there exists $\delta$ such that $|y-z| < \delta$ implies $|D_j g(y) - D_j g(z)| < \varepsilon$. This means that if $|h| < \delta$, $K$ is a compact set that contains the support of $D_j g$, and

$$x - K + B_\delta(0) = \{x - t + z : \ t \in K, \ z \in B_\delta(0)\},$$

$$\begin{aligned}
|N_j(x,h) - f * D_j g(x)| &= \left| \int_{\mathbb{R}^d} f(t)\, D_j g(x - t + \gamma_{t,h} e_j) dt - \int_{\mathbb{R}^d} f(t) D_j g(x-t) dt \right| \\
&= \left| \int_{x-K+B_\delta(0)} f(t) \big[ D_j g(x - t + \gamma_{t,h} e_j) - D_j g(x-t) \big] dt \right| \\
&\le \int_{x-K+B_\delta(0)} |f(t)| \big| D_j g(x - t + \gamma_{t,h} e_j) - D_j g(x-t) \big| dt \\
&\le \int_{x-K+B_\delta(0)} |f(t)|\, \varepsilon\, dt \\
&\le \|f\|_p\, \varepsilon\, m(x - K + B_\delta(0))^{1/q}.
\end{aligned}$$

This shows that $\lim_{h\to 0} N_j(x,h) = f * D_j g(x)$. So $f * g$ is differentiable at $x$ and its partial derivative is $D_j(f * g) = f * D_j g$. For arbitrary derivatives, we can iterate this result as long as $g$ is sufficiently differentiable.

(iii) Assume for simplicity that $g \ge 0$ (nothing really changes if we don't; just a few more absolute values). First, with the change of variable $x = z/n$,

$$\int_{\mathbb{R}^d} g_n(x)\, dx = \int_{\mathbb{R}^d} n^d g(nx)\, dx = \int_{\mathbb{R}^d} g(z)\, dz = 1.$$

Then we have (applying Hölder in the last step, note that $\|g_n^{1/q}\|_q = 1$)

$$\begin{aligned}
|f * g_n(x) - f(x)| &= \left| \int_{\mathbb{R}^d} [f(x-t) - f(x)]\, g_n(t)\, dt \right| \\
&\le \int_{\mathbb{R}^d} |f(x-t) - f(x)|\, g_n^{1/p}\, g_n^{1/q}\, dt \\
&\le \left( \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p\, g_n(t)\, dt \right)^{1/p}.
\end{aligned}$$

Thus (using Tonelli and that $\|g_n\|_1 = 1$)

$$\begin{aligned}\int_{\mathbb{R}^d} |f * g_n(x) - f(x)|^p \, dx &\le \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, g_n(t) \, dt \, dx \\ &= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, dx \right) g_n(t) \, dt \\ &= \int_{\mathbb{R}^d} \|f_t - f\|_p^p \, g_n(t) \, dt.\end{aligned}$$

By $L^p$ continuity (Lemma 2.8.19), we have

$$\lim_{|t| \to 0} \|f_t - f\|_p = 0,$$

where $f_t(x) = f(x-t)$. So, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|t| < \delta$ implies $\|f_t - f\|_p < \varepsilon^{1/p}$. Then

$$\int_{|t|<\delta} \|f_t - f\|_p^p \, g_n(t) \, dt \le \varepsilon \|g_n\|_1 = \varepsilon. \tag{2.51}$$

We also have, for any $t$,

$$\begin{aligned}\int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, dx &\le \int_{\mathbb{R}^d} (|f(x-t)| + |f(x)|)^p \, dx \\ &\le 2^p \int_{\mathbb{R}^d} (|f(x-t)|^p + |f(x)|^p) \, dx \\ &= 2^{p+1} \|f\|_p^p.\end{aligned}$$

Since

$$\int_{|t| \le n\delta} g(t) \, dt \to 1$$

by Monotone Convergence, using again the change of variable $x = z/n$

$$\int_{|t| \ge \delta} g_n(t) \, dt = \int_{|t| \ge \delta} n^d \, g(nt) \, dt = \int_{|t| \ge n\delta} g(t) \, dt \xrightarrow[n \to \infty]{} 0.$$

We obtain

$$\int_{|t| \ge \delta} \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, dx \, g_n(t) dt \le 2^{p+1} \|f\|_p^p \int_{|t| \ge \delta} g_n \xrightarrow[n \to \infty]{} 0.$$

Combining this last equation with (2.51),

$$\limsup_{n \to \infty} \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, dx \right) g_n(t) \, dt \le \varepsilon.$$

By the Limsup Routine, $\|f * g_n - f\|_p \to 0$.

□

*2.8.23. Remark* . We will prove in a while (Corollary 2.11.10) that in the situation of Proposition 2.8.22, if the function $g$ is also in $L^\infty$, then $f * g_n(x) \to f(x)$ almost everywhere.

To appreciate the next result, it is worth emphasizing that it is not entirely obvious how to produce a nonzero infinitely differentiable function with compact support, let alone enough of them to be dense in $L^p(\mathbb{R}^d)$ (the issue being "compact support"; the reader has surely seen plenty of infinitely differentiable functions since calculus, but none of them had compact support). It is also not true in general that $f * g$ has compact support if $g$ does. For example we can take any non-negative integrable function $f$ with support $\mathbb{R}$, and $g = 1_{[0,1]}$; then

$$(f * g)(x) = \int_{\mathbb{R}} f(t)\, 1_{[0,1]}(x-t)\, dt = \int_{x-1}^{x} f(t)\, dt$$

is nonzero for all $x \in \mathbb{R}$. What is true is that if *both* $f, g$ have compact support, then $f * g$ has compact support (Exercise 2.8.16). Proposition 2.8.24 brings in the last two ingredients to approximate functions in $L^p$ with smooth functions.

***2.8.24. Proposition.***

*Let $p \in [1, \infty)$. The set $C_c^\infty(\mathbb{R}^d)$ of infinitely differentiable functions with compact support is dense in $L^p(\mathbb{R}^d)$.*

*Proof.* Let

$$h_1(x) = \begin{cases} e^{-1/x^2}, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{2.52}$$

Then $h_1 \in C^\infty(\mathbb{R})$. Now, given any interval $[a, b]$, form

$$h_{a,b}(x) = h_1(x-a)h_1(b-x).$$

Then $h_{a,b} \in C^\infty(\mathbb{R})$, with support $[a, b]$. Given any box $B = \prod_{j=1}^d [a_j, b_j] \subset \mathbb{R}^d$, the function

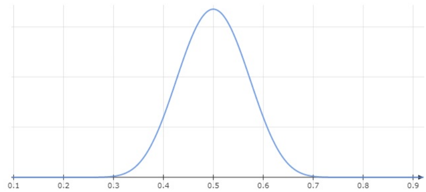
$$g(x) = h_{a_1,b_1}(x_1) \cdots h_{a_d,b_d}(x_d)$$

is nonzero, $C^\infty$, and with support equal to $B$. We can normalize $g$ so that $\|g\|_1 = 1$. By Proposition 2.8.22, we have $\|f - f * g_n\|_p \to 0$, and $f * g_n \in C^\infty(\mathbb{R}^d)$. We are not guaranteed that $f * g_n$ has compact support, though. But by Proposition 2.8.18 there exist bounded functions $f_n \in L^p(\mathbb{R}^d)$ with compact support and such that $\|f - f_n\|_p \to 0$. Now by Proposition 2.8.22 and Exercise 2.8.16 we have $f_n * g_n \in C_c^\infty(\mathbb{R}^d)$ and

$$\begin{aligned} \|f - f_n * g_n\|_p &\le \|f - f * g_n\|_p + \|f * g_n - f_n * g_n\|_p \\ &\le \|f - f * g_n\|_p + \|f - f_n\|_p \, \|g_n\|_1 \to 0. \end{aligned}$$

$\square$

Functions like $g$ in the proof of Proposition 2.8.24 (compactly supported, non-negative, $C^\infty$) are often called **mollifiers**. Below is a picture of $1000h_{0,1}(x)$. The maximum is $1000/e^8 \simeq 0.336$.

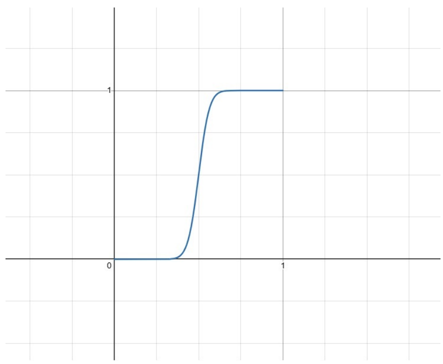



It is remarkable that any single compactly supported $g$ can be used to generate the family $g_n$ in Proposition 2.8.22. When $p = 1$, the inequality $\|f * g\|_1 \le \|f\|_1\|g\|_1$ shows that the convolution can be used to define a product in $L^1$, which becomes an algebra. In terms to be discussed much later in Chapter 9, $L^1(\mathbb{R}^d)$ is a Banach algebra with an approximate unit.

On the real line, if desired, one can get the mollifier to look more like a step. Concretely, one can form

$$g_1(x) = \frac{h_1(x)}{h_1(x) + h_1(1-x)}.$$

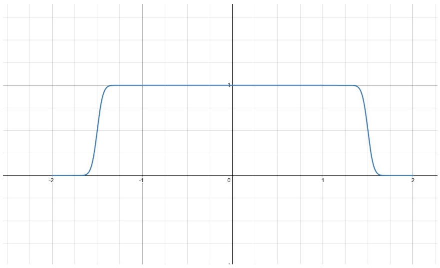
This function is in $C^\infty(\mathbb{R})$. Since $0 \le g_1(x) \le h_1(x)$ for all $x$, we get that $g_1^{(k)}(0) = 0$ for all $k$ (Exercise 2.8.30). Looking at $0 \le 1 - g_1(x) \le h_1(1-x)$ we see that $g_1^{(k)}(1) = 0$ also for all $k$ (again by Exercise 2.8.30). Here is what $g_1$ looks like:




And with this we can form

$$k_1(t) = \begin{cases} g_1(t+2), & t < -1 \\ 1, & |t| \le 1 \\ 1 - g(t-1), & t > 1 \end{cases}$$

which looks like this:




Of course we could have used any interval instead of $[-1, 1]$ for the constant part. We will meet the $L^p$ spaces again in Section 5.6.

**2.8.1. Exercises.**

(2.8.1) Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that the relation defined in (2.46) is an equivalence relation. Show that addition, multiplication, and $p$-norms of classes are well-defined by means of their representatives.

(2.8.2) Let $p \in [1, \infty]$ and $f \in L^p(X)$. Show that $\{|f| = \infty\}$ is a nullset.

(2.8.3) Show that $|f| \leq \|f\|_\infty$ a.e.

(2.8.4) Show that if $(X, \mathcal{A}, \mu)$ is a measure space and $\{f_n\} \subset L^p(X)$, then

$$\Big\| \sum_{k=1}^\infty |f_k| \Big\|_p \leq \sum_{k=1}^\infty \|f_k\|_p,$$

and the inequality still holds even if one or both sides are infinite.

(2.8.5) Let $p \in [1, \infty)$. Show that if $a \in \ell^p(\mathbb{N})$, then $a$ is bounded. Can we say the same for $f \in L^p(\mathbb{R})$?

(2.8.6) Show that the hypothesis that $\|f\|_r < \infty$ for some $r$ in Proposition 2.8.11 cannot be omitted.

(2.8.7) Show that a Cauchy sequence $\{f_n\} \subset L^p(X)$ is uniformly bounded in $L^p(X)$: that is, there exists $c > 0$ such that $\|f_n\|_p < c$ for all $n$.

(2.8.8) Show that $\ell^p(\mathbb{N})$, $1 \leq p < \infty$ is separable, while $\ell^\infty(\mathbb{N})$ is not separable.

(2.8.9) Show that the subspace

$$c_{00} = \{x : \mathbb{N} \to \mathbb{C}, \text{ with finite support}\}$$

is dense in $\ell^p(\mathbb{N})$, $1 \leq p < \infty$. What about the case $p = \infty$?

(2.8.10) Show that if $f \in L^1(\mu)$, then for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\displaystyle\int_E |f| d\mu < \varepsilon$ whenever $\mu(E) < \delta$. Can $\delta$ be chosen independently of $f$?

(2.8.11) Let $(X, \mathcal{A}, \mu)$ be a measure space, $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, and $f \in L^p(X)$, $g \in L^q(X)$. Show that the following statements are equivalent:

(i) $\|fg\|_1 = \|f\|_p \|g\|_q$ ;

(ii) either $f = 0$, $g = 0$, or there exists $\alpha \in (0, \infty)$ with $|f|^p = \alpha |g|^q$.

*(Hint: for the nontrivial part, try to undo the proof of Young's Inequality)*

(2.8.12) Let $(X, \mathcal{A}, \mu)$ be a measure space and $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in L^p(X)$, $g \in L^q(X)$. Show that the following statements are equivalent:

(i) $\|f + g\|_p = \|f\|_p + \|g\|_p$;

(ii) there exist $\alpha, \beta \geq 0$, not both zero, with $\alpha f = \beta g$ a.e.

(2.8.13) Show by example that the $p$-norm is not a norm when $p < 1$.

(2.8.14) Prove the following generalization of Hölder's Inequality for functions $f_1, \ldots, f_n$. Namely, if $p_1, \ldots, p_n \geq 1$ with $\sum_{j=1}^n \frac{1}{p_j} = 1$, and $f_j \in L^{p_j}(X)$, $j = 1, \ldots, n$. Then

$$\|f_1 \cdots f_n\|_1 \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}. \tag{2.53}$$

(2.8.15) Use Hölder's Inequality to prove the following more general inequality: if $f \in L^p(X)$, $g \in L^q(X)$, where $p, q \geq 1$ and $r$ is such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, then

$$\|fg\|_r \leq \|f\|_p \|g\|_q. \tag{2.54}$$

(2.8.16) Show that if $f, g$ have compact support, then $f * g$ has compact support.

(2.8.17) *(Young's Convolution Inequality)* Prove Young's Convolution Inequality (2.50). *(Hint: for non-negative $f, g$ and $p, q, r$ finite,*

$$(f * g)(x) = \int_{\mathbb{R}^d} f(t)^{p/r} g(x-t)^{q/r} \, f(t)^{p/p_1} \, g(x-t)^{q/p_2} \, dt;$$

*then use (2.53) for the exponents $r, p_1, p_2$, where $\frac{1}{p_1} = \frac{1}{p} - \frac{1}{r}$ and $\frac{1}{p_2} = \frac{1}{q} - \frac{1}{r}$)*

(2.8.18) *(Interpolation)*. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $r, s \in [1, \infty]$ with $r < s$. Show that if $f \in L^r(X) \cap L^s(X)$, then $f \in L^p(X)$ for all $p \in (r, s)$.

(2.8.19) Let $f_n : X \to [0, \infty)$ be measurable for all $n$, with $f_1 \geq f_2 \geq \cdots \geq 0$, with $f_n \to f$ pointwise, and such that there exists $n_0$ with $f_{n_0} \in L^1(\mu)$. Prove that

$$\lim_{n \to \infty} \int_X f_n d\mu = \int_X f d\mu.$$

Show that the assertion fails if the $L^1$ condition is omitted.

(2.8.20) Let $X$ be a set. Show that if $1 \leq p < q < \infty$, then $\ell^q(X) \subset \ell^p(X) \subset \ell^\infty(X)$.

(2.8.21) For some measures, $1 \leq r < s$ implies $L^r(\mu) \subset L^s(\mu)$; for others, the reverse inclusion holds; for others, $L^r(\mu) = L^s(\mu)$; and still for others, no inclusion holds if $r \neq s$. Show examples of all these situations, and find conditions on $\mu$ under which each case occurs.

(2.8.22) Given $p \geq 1$, find $f \in L^p(\mathbb{R})$ such that $f \notin L^q(\mathbb{R})$ for any $q \neq p$.

(2.8.23) Do Propositions 2.8.14, 2.8.16 and 2.8.18 hold in $\ell^\infty(\mathbb{N})$? Give proofs or counterexamples.

(2.8.24) Prove Lemma 2.8.19.

(2.8.25) Show that (2.48) can fail when $p=\infty$, even if $\mu(X)<\infty$.

(2.8.26) Let $(X,\mu)$ and $(Y,\nu)$ be $\sigma$-finite complete measure spaces such that $L^2(X)$ and $L^2(Y)$ are separable. Fix orthonormal bases $\{f_n\}$ and $\{g_n\}$ for $L^2(X)$ and $L^2(Y)$ respectively. Show that $\{f_n(x)g_m(y)\}_{n,m}$ is an orthonormal basis for $L^2(X\times Y)$.

(2.8.27) Let $(X,\Sigma,\mu)$ be a measure space. A sequence $\{f_n\}$ of complex measurable functions on $X$ is said to *converge in measure* to the measurable function $f$ if for every $\varepsilon>0$ there exists $N$ such that

$$\mu(\{x:\ |f_n(x)-f(x)|>\varepsilon\})<\varepsilon,\quad n>N.$$

Prove:

(i) If $\mu(X)<\infty$ and $f_n\to f$ a.e., then $f_n\to f$ in measure (*Hint: use Egorov's Theorem*).

(ii) For $1\le p\le\infty$, if $f_n\in L^p(\mu)$ for all $n$ and $\|f_n-f\|_p\to 0$, then $f_n\to f$ in measure.

(iii) If $f_n\to f$ in measure, then there exists a subsequence $f_{n_k}$ that converges to $f$ a.e.

(2.8.28) Let $\{a_n\}_{n\in\mathbb{N}}\subset(0,\infty)$, $(X,\Sigma,\mu)$ a finite measure space, and $f_n:X\to[0,\infty)$, $n\in\mathbb{N}$, measurable functions such that for all $n\in\mathbb{N}$

$$\int_X f_n\,d\mu=n,\qquad\qquad \int_X f_n^2\,d\mu=a_nn^2. \tag{2.55}$$

(i) Show that if the sequence $\{a_n\}$ is bounded, then $f_n$ does not converge to 0 a.e. (*Hint: one possible approach uses Egorov's Theorem*)

(ii) Does the above hold when $\{a_n\}$ is unbounded?

(2.8.29) Suppose that $\mu(X)<\infty$, $f\in L^\infty(\mu)$, $\|f\|_\infty>0$, and

$$\alpha_n=\int_X|f|^n\,d\mu,\qquad n\in\mathbb{N}.$$

Prove that

$$\lim_{n\to\infty}\frac{\alpha_{n+1}}{\alpha_n}=\|f\|_\infty.$$

(2.8.30) Let $h_1$ as in (2.52). Suppose that $g\in C^\infty(-\delta,\delta)$ for some $\delta>0$ and $0\le g(x)\le h_1(x)$ for all $x\in(0,\delta)$. Show that $g^{(k)}(0)=0$ for all $k\in\mathbb{N}$.

(2.8.31) *(this is not an easy one; the topic of which $L^p$ spaces are separable is subtle)* Let $(X,\mathcal{A},\mu)$ be a $\sigma$-finite measure space such that $\mathcal{A}$ is countably generated. Show that if $\mu$ is $\sigma$-finite, then $L^p(\mu)$ is separable for $p\in[1,\infty)$.

## 2.9. The Riesz–Markov Theorem

The Riesz–Markov Theorem is a result about dual spaces, so it partially belongs in Sections 5.5 and 5.6. But the statement is about measure theory and it does have measure-theoretical implications, hence its presence here.

We will take advantage of the following useful notation from Rudin.

***Notation.*** *If $X$ is a topological space, $K, V \subset X$ with $K$ compact and $V$ open, and $f : X \to \mathbb{C}$, we write $K \prec f$ to mean that $f$ is continuous, $0 \le f \le 1$, and $f|_K = 1$. And we write $f \prec V$ to mean that $f$ is continuous, $0 \le f \le 1$, $\operatorname{supp} f$ is compact, and $\operatorname{supp} f \subset V$.*

In terms of this notation, Urysohn's Lemma says that if $X$ is locally compact Hausdorff, then for any inclusion $K \subset V \subset X$ with $K$ compact and $V$ open, there exists $f$ with $K \prec f \prec V$.

***2.9.1. Proposition.*** ***(Partitions of Unity)***

*Let $X$ be a locally compact Hausdorff space, $K \subset X$ compact, and open sets $V_1, \ldots, V_n \subset X$ with $K \subset V_1 \cup \cdots \cup V_n$. Then there exist $h_1, \ldots, h_n : X \to \mathbb{R}$, with $h_j \prec V_j$ and $K \prec \sum_j h_j$.*

*Proof.* For each $x \in K$ there exists $j(x)$ such that $x \in V_{j(x)}$. By local compactness there exists $W_x$, open, with compact closure and $x \in W_x \subset \overline{W_x} \subset V_{j(x)}$. By compactness of $K$ there exist $x_1, \ldots, x_m$ such that $K \subset \bigcup_{j=1}^{m} W_{x_j}$. Define

$$H_j = \bigcup_{\ell \in I_j} \overline{W_\ell}, \qquad I_j = \{\ell : \ \overline{W_{x_\ell}} \subset V_j\}, \qquad j = 1, \ldots, n.$$

By Urysohn's Lemma there exist functions $g_1, \ldots, g_n$ with $H_j \prec g_j \prec V_j$. Then we define

$$h_1 = g_1, \quad h_2 = (1 - g_1)g_2, \quad \ldots \quad h_n = (1 - g_1) \cdots (1 - g_{n-1})g_n.$$

Since $g_j \prec V_j$, we get $h_j \prec V_j$ for all $j$, as $0 \le h_j \le 1$ and $\operatorname{supp} h_j \subset \operatorname{supp} g_j$. Inductively, suppose that $h_1 + \cdots + h_k = 1 - (1 - g_1) \cdots (1 - g_k)$. Then

$$h_1 + \cdots + h_{k+1} = 1 - (1 - g_1) \cdots (1 - g_k) + (1 - g_1) \cdots (1 - g_k) g_{k+1}$$

$$= 1 - (1 - g_1) \cdots (1 - g_{k+1}).$$

Thus $h_1 + \cdots + h_n = 1 - (1 - g_1) \cdots (1 - g_n)$. If $x \in K$, then there exists $j$ with $x \in H_j$; this implies that $g_j(x) = 1$, and so $\sum_j h_j(x) = 1$. □

A key step in the path to abstraction, both in Measure Theory and in Functional Analysis, is to recognize that an integral is nothing but a way to assign numbers to functions in a linear and positive way. This is made explicit now. The proof—that follows fairly closely the one in [Rud87]—is a tour de force that deserves to be admired.

***2.9.2. Theorem. (Riesz–Markov)***

*Let $T$ be a locally compact Hausdorff space, and $\varphi : C_c(T) \to \mathbb{C}$ linear and positive (meaning that $\varphi(f) \geq 0$ if $f \geq 0$). Then there exists a $\sigma$-algebra $\mathcal{M}$ with $\mathcal{B}(T) \subset \mathcal{M}$ and a measure $\mu$, unique on $\mathcal{B}(T)$, such that*

$$\varphi(f) = \int_T f \, d\mu, \quad f \in C_c(T), \tag{2.56}$$

*and*

*(i) $\mu(K) < \infty$ for all $K$ compact;*

*(ii) $\mu$ is outer regular;*

*(iii) $\mu$ is inner regular on open and finite-measure sets;*

*(iv) $\mu$ is complete.*

*Proof.* Let us consider the uniqueness first. The properties (i)–(iii) guarantee that $\mu$ is determined by its values on compact subsets. Suppose that $\mu_1, \mu_2$ satisfy (2.56) and (i)–(iii) for some $\sigma$-algebra $\mathcal{M}$ that contains the Borel sets. Fix $K \subset T$ compact, and $\varepsilon > 0$. By hypothesis $\mu_2(K) < \infty$ and there exists $V$ open with $K \subset V$ and $\mu_2(V) < \mu_2(K) + \varepsilon$. Using Urysohn's Lemma we construct $f \in C_c(T)$ with

$$\mu_1(K) = \int_T 1_K \, d\mu_1 \leq \int_T f \, d\mu_1 = \varphi(f) = \int_T f \, d\mu_2$$
$$\leq \int_T 1_V \, d\mu_2 = \mu_2(V) \leq \mu_2(K) + \varepsilon.$$

With $\varepsilon$ arbitrary, we conclude that $\mu_1(K) \leq \mu_2(K)$; and, as the roles can be exchanged, $\mu_1(K) = \mu_2(K)$. The inner regularity allows us to pass to open sets, and the outer regularity then implies that $\mu_1 = \mu_2$ on $\mathcal{M}$ (in particular, on all Borel sets).

So now we focus on the existence. We will first construct $\mathcal{M}$ and $\mu$, and then prove their properties through a series of claims. The spirit of what needs to happen is clear: we want $\mu(E) = \varphi(1_E)$; the problem is of course that $1_E$ will not usually be a continuous function, which means that we cannot apply $\varphi$ to it. So we think of approximating $1_V$ from below by continuous functions, and we define

$$\mu(V) = \sup\{\varphi(f) : \ f \prec V\}, \qquad\qquad V \ \text{ open.}$$

And then, as we did with the Lebesgue outer measure a long time ago, we define

$$\mu(E) = \inf\{\mu(V) : \ E \subset V, \ V \ \text{ open}\}, \qquad\qquad E \subset T.$$

To define $\mathcal{M}$, we first consider

$$\mathcal{M}_F = \big\{E \subset T : \mu(E) < \infty \ \text{ and } \mu(E) = \sup\{\mu(K) : K \subset E, \ K \ \text{ compact}\}\big\}.$$

This feels like cheating, because we are baking the regularity in; but we will somehow show that enough sets are in $\mathcal{M}_F$. Finally we define

$$\mathcal{M} = \{E \subset T : \ E \cap K \in \mathcal{M}_F \ \text{ for all } K \text{ compact}\}.$$

Because $\mu$ is defined as an infimum, we have that if $E_1 \subset E_2$ then $\mu(E_1) \leq \mu(E_2)$. Now if $E \subset N$ with $N \in \mathcal{M}$ and $\mu(N) = 0$, then obviously $\mu(E) = 0$; if $K \subset E$ then $\mu(K) = 0$, and so $E \in \mathcal{M}_F$. As the same reasoning applies to $E \cap K$, we get that $E \in \mathcal{M}$, showing that $\mu$ is complete.

Now we go into the first of many steps, where we will build $\mu$ and $\mathcal{M}$ step by step. We state each step and prove it right away.

***I.*** *If $E_1, E_2, \ldots \subset T$, then $\mu\big(\bigcup_n E_n\big) \leq \sum_n \mu(E_n)$.*

We begin with two open sets. If $V_1, V_2$ are open, consider $g \prec V_1 \cup V_2$. By Proposition 2.9.1 there exist $h_1, h_2 \in C_c(T)$ with $h_1 \prec V_1$, $h_2 \prec V_2$, and $h_1 + h_2 = 1$ on $\operatorname{supp} g$. Then $h_1 g \prec V_1$, $h_2 g \prec V_2$, $g = gh_1 + gh_2$, and

$$\varphi(g) = \varphi(h_1 g) + \varphi(h_2 g) \leq \mu(V_1) + \mu(V_2).$$

As this can be done for any $g \prec V_1 \cup V_2$, we have shown that $\mu(V_1 \cup V_2) \leq \mu(V_1) + \mu(V_2)$. Inductively, we can extend this to any finite family of open sets.

For the general case, if there exists $E_k$ with $\mu(E_k) = \infty$, then the inequality holds. So assume that $\mu(E_n) < \infty$ for all $n$. Fix $\varepsilon > 0$. By definition of $\mu$, for each $n$ there exists $V_n$ open with $E_n \subset V_n$ and $\mu(V_n) < \mu(E_n) + \varepsilon/2^n$. Put $V = \bigcup_n V_n$, and fix $f \prec V$. By the compactness of $\operatorname{supp} f$, there exists $n$ such that $\operatorname{supp} f \subset V_1 \cup \cdots \cup V_n$, so $f \prec V_1 \cup \cdots \cup V_n$. Then

$$\varphi(f) \leq \mu(V_1 \cup \cdots \cup V_n) \leq \sum_{k=1}^{n} \mu(V_j) \leq \sum_{k=1}^{n} \mu(E_n) + \varepsilon \leq \sum_{k=1}^{\infty} \mu(E_n) + \varepsilon.$$

Since the above works for any $f \prec V$, we get that $\mu(V) \leq \sum_{k=1}^{n} \mu(E_n) + \varepsilon$. Then, as $\bigcup_n E_n \subset V$,

$$\mu\Big(\bigcup_n E_n\Big) \leq \mu(V) \leq \sum_{n=1}^{\infty} \mu(E_n) + \varepsilon.$$

Finally, as the $\varepsilon$ was arbitrary, we may remove it, and the inequality still stands.

***II.*** *If $K$ is compact, then $K \in \mathcal{M}_F$ and $\mu(K) = \inf\{\varphi(f) : \ K \prec f\}$.*

Fix $f$ with $K \prec f$ and $\alpha \in (0,1)$. Let $V_\alpha = f^{-1}(\alpha, \infty)$, which is open by the continuity of $f$. Then $K \subset V_\alpha$, and $\alpha g \leq f$ for all $g \prec V_\alpha$. Then, as $\varphi$ is positive we have from $0 \leq f - \alpha g$ that $0 \leq \varphi(f - \alpha g) = \varphi(f) - \alpha\varphi(g)$. Thus

$$\mu(K) \leq \mu(V_\alpha) = \sup\{\varphi(g) : \ g \prec V_\alpha\} \leq \frac{1}{\alpha}\,\varphi(f).$$

As $\alpha$ can be chosen arbitrarily close to 1, we get that $\mu(K) \leq \varphi(f) < \infty$ and $K \in \mathcal{M}_F$.

Given $\varepsilon > 0$, there exists $V$ open with $K \subset V$ and $\mu(V) < \mu(K) + \varepsilon$. By Urysohn's Lemma there exists $f$ with $K \prec f \prec V$. Then $\varphi(f) \leq \mu(V) < \mu(K) + \varepsilon$. As this can be done for any $\varepsilon > 0$, we get that

$$\mu(K) = \inf\{\varphi(f) : \ K \prec f\}.$$

***III.*** *For any open set $V \subset T$, $\mu(V) = \sup\{\mu(K) : K \subset V, \ ! \text{ compact}\}$. In particular if $V$ is open and $\mu(V) < \infty$ then $V \in \mathcal{M}_F$.*

Fix $\alpha \in \mathbb{R}$ with $\alpha < \mu(V)$. By definition of $\mu(V)$, there exists $f \prec V$ with $\alpha < \varphi(f)$. Let $K = \operatorname{supp} f \subset V$. If $W$ is open with $K \subset W$, then $f \prec W$ and $\varphi(f) \leq \mu(W)$; as this can be done for any such $W$, we get that

$$\varphi(f) \leq \inf\{\mu(W) : \ K \subset W, \ W \ \text{ open}\} = \mu(K).$$

So $\alpha < \mu(K) < \mu(V)$, showing that we can obtain $\mu(K)$ arbitrarily close to $\mu(V)$ (or arbitrarily large if $\mu(V) = \infty$).

***IV.*** *If $E = \bigcup_n E_n$, with $E_n \in \mathcal{M}_F$ for all $n$ and parwise disjoint, then $\mu(E) = \sum_n \mu(E_n)$. If in addition $\mu(E) < \infty$, then $E \in \mathcal{M}_F$.*

Fix $K_1, K_2$ compact with $K_1 \cap K_2 = \varnothing$. Fix $\varepsilon > 0$. By Urysohn's Lemma there exists $f \in C_c(T)$ with $K_1 \prec f \prec K_2^c$. By ***II*** there exists $g$ with

$K_1 \cup K_2 \prec g$ and $\varphi(g) < \mu(K_1 \cup K_2) + \varepsilon$. It is trivial to check that $K_1 \prec fg$ and $K_2 \prec (1-f)g$. Using ***II*** again,

$$\mu(K_1) + \mu(K_2) \leq \varphi(fg) + \varphi((1-f)g) = \varphi(g) < \mu(K_1 \cup K_2) + \varepsilon.$$

Thus $\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2)$, as the reverse inequality was proven in ***I***. Inductively, we obtain additivity for finitely many pairwise disjoint compact subsets.

If $\mu(E) = \infty$, then already ***I*** gives the equality; so we assume $\mu(E) < \infty$. Fix $\varepsilon > 0$. Since $E_n \in \mathcal{M}_F$ for each $n$, there exists $K_n$ compact with $K_n \subset E_n$ and $\mu(E_n) < \mu(K_n) + \varepsilon/2^n$. Then, using what we just proved (and since the compacts $K_n$ are pairwise disjoint as the $E_n$ are)

$$\mu(E) \geq \mu(K_1 \cup \cdots \cup K_m) = \sum_{n=1}^{m} \mu(K_n) > \sum_{n=1}^{m} \mu(E_n) - \varepsilon.$$

Making $\varepsilon \to 0$, $m \to \infty$ and using the reverse inequality from ***I***, we have shown that $\mu(E) = \sum_n \mu(E_n)$.

To show the last statement, if $\mu(E) < \infty$ then given $\varepsilon > 0$ there exists $m$ with

$$\mu(E) \leq \sum_{n=1}^{m} \mu(E_n) + \varepsilon \leq \sum_{n=1}^{m} \mu(K_n) + 2\varepsilon = \mu\Big(\bigcup_{n=1}^{m} K_n\Big) + 2\varepsilon,$$

which shows that $\mu(E) = \sup\{\mu(K) : \ K \subset E, \ K \ \text{ compact}\}$. Thus $E \in \mathcal{M}_F$.

***V.*** *If $E \in \mathcal{M}_F$ and $\varepsilon > 0$, there exist $K$ compact, $V$ open with $K \subset E \subset V$ and $\mu(V \setminus K) < \varepsilon$.*

By definition of $\mathcal{M}_F$ and $\mu$ there exist $K$ compact and $V$ open with $K \subset E \subset V$ and

$$\mu(V) - \frac{\varepsilon}{2} < \mu(E) < \mu(K) + \frac{\varepsilon}{2}.$$

Since $V \setminus K$ is open and $\mu(V \setminus K) \leq \mu(V) < \mu(E) + \frac{\varepsilon}{2}$, we have that $V \setminus K \in \mathcal{M}_F$ by ***III***. Now using ***IV***, $\mu(K) + \mu(V \setminus K) = \mu(V) < \mu(K) + \varepsilon$, implying that $\mu(V \setminus K) < \varepsilon$.

***VI.*** *If $A, B \in \mathcal{M}_F$, then $A \setminus B$, $A \cup B$, $A \cap B \in \mathcal{M}_F$.*

Fix $\varepsilon > 0$. By ***V***, there exist $K_1, K_2$ compact and $V_1, V_2$ open with $K_1 \subset A \subset V_1$, $K_2 \subset B \subset V_2$, $\mu(V_j \setminus K_j) < \varepsilon$, $j = 1, 2$.

The inclusion

$$A \setminus B \subset V_1 \setminus K_2 \subset (V_1 \setminus K_1) \cup (K_1 \setminus V_2) \cup (V_2 \setminus K_2)$$

implies

$$\mu(A \setminus B) \leq 2\varepsilon + \mu(K_1 \setminus V_2).$$

As $K_1 \setminus V_2 \subset A \setminus B$, and compact, we have shown that $\mu(A \setminus B) = \sup\{\mu(K) : K \subset A \setminus B,\ K \text{ compact}\}$. Thus $A \setminus B \in \mathcal{M}_F$. Now $A \cup B = B \cup (A \setminus B) \in \mathcal{M}_F$ by ***IV***, and $A \cap B = A \setminus (A \setminus B) \in \mathcal{M}_F$.

> ***VII.*** *$\mathcal{M}$ is a $\sigma$-algebra, and $\mathcal{B}(T) \subset \mathcal{M}$.*

Let $A \in \mathcal{M}$, and $K \subset T$ compact. We have $K \in \mathcal{M}_F$ by ***II***, and $A \cap K \in \mathcal{M}_F$ by definition of $\mathcal{M}$. Then $A^c \cap K = K \setminus (A \cap K) \in \mathcal{M}_F$ by ***VI***. As $K$ was arbitrary, $A^c \in \mathcal{M}$. So $\mathcal{M}$ contains complements. We can argue directly that $\varnothing \in \mathcal{M}$, or that $T \in \mathcal{M}$ by ***II*** since $T \cap K = K$; in either case, both $\varnothing$ and $T$ are in $\mathcal{M}$.

If $\{A_n\} \subset \mathcal{M}$, fix again a compact $K \subset T$. Define sets $B_1 = A_1 \cap K$, $B_n = (A_n \cap K) \setminus \bigcup_{k=1}^{n-1} B_k$. These sets are pairwise disjoint by construction, $B_1 \in \mathcal{M}_F$ by definition of $\mathcal{M}$, and $B_n \in \mathcal{M}_F$ by ***VI***. Write $A = \bigcup_n A_n$. As $\mu(A \cap K) \le \mu(K) < \infty$ by ***II***, we have $A \cap K = \bigcup_k B_k \in \mathcal{M}_F$ by ***IV***, and then $A \in \mathcal{M}$, showing that $\mathcal{M}$ is a $\sigma$-algebra.

If $A$ is closed, then $A \cap K$ is compact, so $A \cap K \in \mathcal{M}_F$, and $A \in \mathcal{M}$; as $\mathcal{M}$ is a $\sigma$-algebra that contains the closed sets, $\mathcal{B}(T) \subset \mathcal{M}$.

> ***VIII.*** *$\mathcal{M}_F = \{E \in \mathcal{M} : \ \mu(E) < \infty\}$. This is (iii).*

Given $E \in \mathcal{M}_F$, for any $K$ compact $E \cap K \in \mathcal{M}_F$ by ***II*** and ***VI***, so $E \in \mathcal{M}$.

Conversely, if $E \in \mathcal{M}$ with $\mu(E) < \infty$, fix $\varepsilon > 0$. By definition of $\mu$ there exists $V$ open, $E \subset V$, $\mu(V) < \infty$. By ***III*** we have that $V \in \mathcal{M}_F$ and using ***V*** on the set $V$ we have that there exists $K$ compact with $K \subset V$ and $\mu(V \setminus K) < \varepsilon$. Since $E \cap K \in \mathcal{M}_F$ by definition of $\mathcal{M}$, there exists $H \subset E \cap K$, compact, with $\mu(E \cap K) < \mu(H) + \varepsilon$. Using that $E \subset (E \cap K) \cup (V \setminus K)$, and ***I***,

$$\mu(E) \le \mu(E \cap K) + \mu(V \setminus K) < \mu(H) + 2\varepsilon.$$

This shows that $\mu(E) = \sup\{\mu(H) : \ H \subset E,\ H \text{ compact}\}$, giving us that $E \in \mathcal{M}_F$.

> ***IX.*** *$\mu$ is a measure on $\mathcal{M}_F$.*

Let $\{E_k\} \subset \mathcal{M}$, pairwise disjoint, and let $E = \bigcup_n E_n$. By ***I*** we have that $\mu(E) \le \sum_n \mu(E_n)$. If $\mu(E_n) = \infty$ for some $n$, then the equality holds trivially. If $\mu(E) = \infty$, the equality also holds trivially. So we may assume that $\mu(E_n) < \infty$ for all $n$, and that $\mu(E) < \infty$. This gives us $E, E_n \in \mathcal{M}_F$ for all $n$ by ***VIII***. By ***IV***, $\mu(E) = \sum_n \mu(E_n)$.

> ***X.*** *$\varphi(f) = \displaystyle\int_T f\,d\mu$ for any $f \in C_c(T)$.*

If $g \in C_c(T)$ is real-valued, then $g = g^+ - g^-$ with $g^+, g^- \geq 0$, $g^+, g^- \in C_c(T)$, and so $\varphi(g) = \varphi(g^+) - \varphi(g^-) \in \mathbb{R}$. That is, $\varphi$ maps real functions to real functions. This lets us assume for the rest of the argument that $f$ is real, because if $f$ is complex then

$$\begin{aligned}\varphi(f) &= \varphi(\operatorname{Re} f + i \operatorname{Im} f) = \varphi(\operatorname{Re} f) + i\varphi(\operatorname{Im} f)\\ &= \int_T \operatorname{Re} f\,d\mu + i\int_T \operatorname{Im} f\,d\mu = \int_T f\,d\mu.\end{aligned}$$

Hence fix $f \in C_c(T)$, real. Let $K = \operatorname{supp} f$. As $f(K)$ is compact, there exist $a, b \in \mathbb{R}$ such that $f(K) \subset [a, b]$. Fix $\varepsilon > 0$ and choose a partition $y_0 < a < y_1 < \cdots < y_n = b$, with $y_k - y_{k-1} < \varepsilon$ for all $k$. Let

$$E_k = K \cap f^{-1}(y_{k-1}, y_k], \qquad k = 1, \ldots, n.$$

As $E_k$ is the intersection of two Borel sets, it is Borel, and so $E_k \in \mathcal{M}_F$ (since $\mu(E_k) \leq \mu(K) < \infty$). By construction, the sets $E_1, \ldots, E_n$ are pairwise disjoint, and $\displaystyle\bigcup_{k=1}^n E_k = K$. Choose open sets $V_1, \ldots, V_n$ with $E_k \subset V_k$ and $\mu(V_k) < \mu(E_k) + \varepsilon/n$. By shrinking $V_k$ is necessary (using the continuity of $f$), we may assume that $f(t) < y_k + \varepsilon$ for all $t \in V_k$. Use Proposition 2.9.1 to get a partition of unity $h_1, \ldots, h_n$ with $h_k \prec V_k$ for all $k$ and $\sum_k h_k = 1$ on $K$. This allows us to write $f = \sum_k h_k f$. Also,

$$\sum_{k=1}^n \mu(E_k) = \mu(K) \leq \varphi(\sum_k h_k) = \sum_k \varphi(h_k)$$

(since $K \prec \sum_k h_k$ and by ***II***). For the computation below, we will use that $h_k f \leq (y_k + \varepsilon) h_k$ (since $h_k \geq 0$ with support inside $V_k$); also, that $\varphi(h_k) \leq \mu(V_k) < \mu(E_k) + \varepsilon/n$ (the first inequality, by definition of $\mu$); and that $y_k - \varepsilon < f$ on $E_k$ . Then, adding $|a|$ so that $|a| + y_k + \varepsilon > 0$,

$$\begin{aligned}\varphi(f) &= \sum_{k=1}^n \varphi(h_k f) \leq \sum_{k=1}^n (y_k + \varepsilon)\varphi(h_k)\\ &= \sum_{k=1}^n (|a| + y_k + \varepsilon)\varphi(h_k) - |a| \sum_{k=1}^n \varphi(h_k)\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=1}^{n}(|a| + y_k + \varepsilon)\big(\mu(E_k) + \frac{\varepsilon}{n}\big) - |a|\,\mu(K)\\
&= \sum_{k=1}^{n}(y_k - \varepsilon)\mu(E_k) + 2\varepsilon\mu(K) + \varepsilon\,|a| + \frac{\varepsilon}{n}\sum_{k=1}^{n}(y_k + \varepsilon)\\
&= \int_T \sum_{k=1}^{n}(y_k - \varepsilon)\,1_{E_k}\,d\mu + 2\varepsilon\mu(K) + \varepsilon\,|a| + \varepsilon(b + \varepsilon)\\
&\leq \int_T f\,d\mu + \varepsilon(2\mu(K) + |a| + b + \varepsilon).
\end{aligned}$$

Being free to choose $\varepsilon$, we have shown that $\varphi(f) \leq \int_T f\,d\mu$ for any $f \in C_c(T)$. As this also works for $-f$ and $\varphi$ is linear, we get (2.56). □

*2.9.3. Remark* . In the Riesz–Markov Theorem, the equality (2.56) implies condition (i) (Exercise 2.9.1). In the case where every open set is $\sigma$-compact (that is, a countable union of compact subsets) the condition (i) implies both (ii) and (iii); see Proposition 2.9.11

*2.9.4. Remark* . The positivity condition in Theorem 2.9.2 may seem a bit contrived. And in a sense it is, as a more general result can be proven. Such result, Theorem 5.6.15, requires a bit more preparation (complex measures, as in Section 2.10, and the dual from Section 5.5), hence its location much further down the text. In the end, though, the positivity condition is very natural and we will see more about those kind of conditions in Chapter 10 and beyond.

*2.9.5. Remark* . The Riesz–Markov Theorem can be used to construct the Lebesgue measure from scratch. Simply take $T = \mathbb{R}$ with its usual topology, and let $\varphi : C_c(\mathbb{R}) \to \mathbb{C}$ be given by

$$\varphi(f) = \int_{-\infty}^{\infty} f(t)\,dt.$$

Since $f$ has compact support, the integral is a regular Riemman integral and not improper; that is, we may as well integrate over any closed interval that contains $\operatorname{supp} f$. As the measure $\mu$ is the unique measure on $\mathcal{M}$—which includes $\mathcal{B}(\mathbb{R})$—satisfying the properties in Riesz–Markov, it is the Lebesgue measure, and the uniqueness also guarantees that $\mathcal{M} = \mathcal{M}(\mathbb{R})$ (see Proposition 2.9.10 below).

The theorem also saves us the pain of repeating all the work done in constructing the Lebesgue measure if we want to do it in $\mathbb{R}^n$ instead of on $\mathbb{R}$, be it "by hand" or as a product measure for $\mathbb{R}^d$, $d > 1$. Simply applying Riesz–Markov to $\varphi : C_c(\mathbb{R}^n) \to \mathbb{C}$ given by

$$\varphi(f) = \int_B f(x)\,dx, \qquad\qquad f \in C_c(\mathbb{R}^n),$$

where $B$ is a box that contains $\operatorname{supp} f$ and the integral is the iterated integral over the box.

*2.9.6. Example*. Let $T$ be any locally compact Hausdorff space, Fix $t_0 \in T$, and define $\varphi : C_c(T) \to \mathbb{C}$ by $\varphi(f) = f(t_0)$. If we consider the measure $\mu$ we obtain from Riesz–Markov, given any $V \subset T$ open we have

$$\mu(V) = \sup\{\varphi(f) : \ f \prec V\} = \sup\{f(t_0) : \ f \prec V\} = \begin{cases} 1, & t_0 \in V \\ 0, & t_0 \notin V \end{cases}$$

so $\mu = \delta_{t_0}$ on open sets. And for any $E \in \mathcal{M}$ we have $\mu(E) = \inf\{\mu(V) : \ E \subset V,\ V \ \text{open}\} = \delta_{t_0}(E)$, showing that $\mu = \delta_{t_0}$ for any $E \subset T$ (recall that in the proof of the Riesz–Markov Theorem, $\mu$ is defined on all subsets). It is then trivial to show that $\mathcal{M}_F = \mathcal{P}(T)$ and so $\mathcal{M} = \mathcal{P}(T)$.

*2.9.7. Remark*. Without local compactness, uniqueness can fail in the Riesz–Markov Theorem. If $T$ is Hausdorff but not locally compact, there exists $t_0 \in T$ such that no open neighbourhood of $t_0$ is contained in a compact set. If $f \in C_c(T)$ and $f \neq 0$, then $V = f^{-1}(\mathbb{C} \setminus \{0\})$ is open and $V \subset \operatorname{supp} f$. This means that $t_0 \notin V$, and hence $f(t_0) = 0$. If we define $\psi(f) = \int_T f\,d\delta_{t_0}$, then $\psi(f) = 0$ for all $f \in C_c(T)$. Thus the zero functional can be represented both by the zero measure and by $\delta_{t_0}$.

*2.9.8. Remark*. The regularity is also crucial for the uniqueness of the measure in Riesz–Markov. We refer to Exercise 2.9.6 for an example of a linear functional that can be represented both by a regular and a non-regular measure.

As long as $T$ is not a terribly bad behaved topological space, the $\sigma$-algebra $\mathcal{M}$ and the measure $\mu$ obtained in the Riesz–Markov Theorem have further nice properties—regarding regularity—than those stated in the theorem. Compare with Proposition 2.3.28.

**2.9.9. Definition.**

*A topological space $T$ is **$\sigma$-compact** if it can be written as a countable union of compact sets.*

***2.9.10. Proposition.***

*Let $T$ be locally compact Hausdorff, $\sigma$-compact, and $(\mathcal{M}, \mu)$ as in the Riesz–Markov Theorem. Then $\mu$ is Radon. More concretely,*

*(i) if $E \in \mathcal{M}$ and $\varepsilon > 0$, there exist $F$ compact, $V$ open with $F \subset E \subset V$ and $\mu(V \setminus F) < \varepsilon$;*

*(ii) $\mu$ is regular by Borel sets;*

*(iii) if $E \in \mathcal{M}$, there exist $A, B \in \mathcal{B}(T)$ with $A$ an $F_\sigma$ and $B$ a $G_\delta$, $A \subset E \subset B$ and $\mu(B \setminus A) = 0$; in particular, $\mu(A) = \mu(E) = \mu(B)$;*

*(iv) $E \in \mathcal{M}$ if and only if there exists $B \in \mathcal{B}(T)$ and $N \in \mathcal{M}$ with $E = B \cup N$ and $\mu(N) = 0$.*

*Proof.* The fact that $T$ is $\sigma$-compact allows us to write $T = \bigcup_n K_n$, an increasing union of compact subsets.

If $E \in \mathcal{M}$, we have $\mu(E \cap K_n) < \infty$ for all $n$. By the regularity of $\mu$ there exists $V_n \subset T$, open, with $E \cap K_n \subset V_n$ and $\mu(V_n \setminus (E \cap K_n)) < \frac{\varepsilon}{2^{n+1}}$. Putting $V = \bigcup_n V_n$, we get

$$V \setminus E = V \cap E^c = \bigcup_n (V_n \cap E^c) \subset \bigcup_n V_n \cap (E \cap K_n)^c = \bigcup_n (V_n \setminus (E \cap K_n)),$$

and so $\mu(V \setminus E) < \frac{\varepsilon}{2}$. As $E^c \in \mathcal{M}$, we can repeat the argument to see that there exists $W$ open with $E^c \subset W$ and $\mu(W \setminus E^c) < \frac{\varepsilon}{2}$. With $F = W^c$, we get that $F$ is closed and

$$E \setminus F = E \cap F^c = F^c \setminus E^c = W \setminus E^c.$$

Thus $\mu(E \setminus F) < \varepsilon/2$ and

$$\mu(V \setminus F) = \mu((V \setminus E) \cup (E \setminus F)) = \mu(V \setminus E) + \mu(E \setminus F) < \varepsilon.$$

For regularity, what is missing is the case of infinite-measure sets. Given any $E \in \mathcal{M}$ and $\varepsilon > 0$, by the above there exists $F$ closed with $F \subset E$ and $\mu(E \setminus F) < \varepsilon$. Writing again $T = \bigcup_n K_n$ as an increasing union of compact subsets, we can write $F = \bigcup_n (F \cap K_n)$. So $E \setminus F = \bigcap_n E \setminus (F \cap K_n)$. By continuity of the measure, there exists $n$ such that $\mu(E \setminus (F \cap K_n)) < \varepsilon$. So we have $F \cap K_n$ compact, $F \cap K_n \subset E$, and $\mu(E \setminus (F \cap K_n)) < \varepsilon$.

By the above, for each $k \in \mathbb{N}$ there exist $F_k$ closed, $V_k$ open with $F_k \subset E \subset V_k$ and $\mu(V_k \setminus F_k) < 1/k$. Let $A = \bigcup_k F_k$, $B = \bigcup_k V_k$. Then $A \subset E \subset B$ and $B \setminus A \subset V_k \setminus F_k$ for all $k$; thus $\mu(B \setminus A) = 0$.

Now using the $F_\sigma$ set $A$ from the previous paragraph, if $E \in \mathcal{M}$ we can write $E = A \cup N$, with $N = E \setminus A$. As $A \in \mathcal{B}(T) \subset \mathcal{M}$, $E \setminus A \subset B \setminus A$ is measurable by being a subset of a nullset on a complete measure space; hence $\mu(N) = 0$. The converse follows directly from the fact that $\mathcal{B}(T) \subset \mathcal{M}$. ☐

A consequence of Proposition 2.9.10 is that the $\sigma$-algebra in Theorem 2.9.2 is unique. Indeed, if we had $\mathcal{M}_1$ and $\mathcal{M}_2$, take $E \in \mathcal{M}_1$. Then, in the notation of Proposition 2.9.10, $E = A \cup (E \setminus A)$. As $A$ is Borel, we have $A \in \mathcal{M}_2$. And $E \setminus A \subset B \setminus A$; as $B \setminus A$ is Borel, it is in $\mathcal{M}_2$. Since $B \setminus A$ is a nullset and $\mathcal{M}_2$ is complete, $E \setminus A \in \mathcal{M}_2$ and hence $E \in \mathcal{M}_2$. The roles can be reversed, hence $\mathcal{M}_1 = \mathcal{M}_2$.

Another consequence of Theorem 2.9.2 is that many measures are regular under fairly mild conditions:

***2.9.11. Proposition.***

*Let $T$ be locally compact Hausdorff, with every open subset $\sigma$-compact. Let $\lambda$ be a Borel measure on $T$ such that $\lambda(K) < \infty$ for all compact $K \subset T$. Then $\lambda$ is Radon.*

*Proof.* Define $\varphi : C_c(T) \to \mathbb{C}$ by

$$\varphi(f) = \int_T f \, d\lambda, \qquad f \in C_c(T).$$

This well-defined because $\lambda(\operatorname{supp} f) < \infty$ for all $f$, by hypothesis. By Riesz–Markov and Proposition 2.9.10 there exists a $\sigma$-algebra $\mathcal{M}$ with $\mathcal{B}(T) \subset \mathcal{M}$ and a Radon measure $\mu$ on $\mathcal{M}$ such that

$$\int_T f \, d\lambda = \int_T f \, d\mu, \qquad f \in C_c(T).$$

Given $V \subset T$ open, by hypothesis $V = \bigcup_n K_n$ with $K_n$ compact for all $n$. We may assume without loss of generality that the sequence $\{K_n\}$ is increasing. By Urysohn's Lemma for each $n$ there exists $f_n \in C_c(T)$ with $K_n \prec f \prec V$. Let $g_k = \max\{f_1, \ldots, f_k\} \in C_c(T)$. Then $g_k \nearrow 1_V$. By

Monotone Convergence,

$$\lambda(V)=\int_T 1_V\,d\lambda=\lim_k\int_T g_k\,d\lambda=\lim_k\int_T g_k\,d\mu=\int_T 1_V\,d\mu=\mu(V).$$

Now if $E$ is Borel, fix $\varepsilon>0$. By Proposition 2.9.10 there exist $F,V\subset T$ with $F$ closed, $V$ open, $F\subset E\subset V$, and $\mu(V\setminus F)<\varepsilon$. As $V\setminus F$ is open, $\lambda(V\setminus F)=\mu(V\setminus F)<\varepsilon$. Then

$$\lambda(E)\le\lambda(V)=\mu(V)\le\mu(F)+\varepsilon\le\mu(E)+\varepsilon,$$
$$\mu(E)\le\mu(V)=\lambda(V)\le\lambda(F)+\varepsilon\le\lambda(E)+\varepsilon.$$

As this can be done for any $\varepsilon>0$, we obtain that $\lambda(E)=\mu(E)$. Note that the estimates still work when $E$ has infinite measure. □

The hypothesis of $\sigma$-compactness in Proposition 2.9.11 cannot be relaxed, as shown by Exercise 2.9.5. So it is not possible to obtain full regularity in Riesz–Markov without additional hypotheses.

**2.9.1. Exercises.**

(2.9.1) Show that under the hypotheses of Riesz–Markov, if a measure $\mu$ satisfies (2.56) then $\mu(K)<\infty$ for all $K$ compact.

(2.9.2) Use Riesz–Markov to construct Lebesgue measure in $\mathbb{R}^n$, $n\in\mathbb{N}$. Prove that the measure you constructed is **the** Lebesgue measure, by showing that it agrees with the Lebesgue outer measure on boxes. Show that the $\sigma$-algebra $\mathcal{M}$ from the theorem is $\mathcal{M}(\mathbb{R}^n)$.

(2.9.3) Let $X$ be a topological space and $\mu$ a Borel measure. The **support** of $\mu$ is the set

$$\operatorname{supp}\mu=\{x\in X:\ \mu(V)>0\ \text{ for all } V \text{ open with } x\in V\}.$$

Prove that $\operatorname{supp}\mu$ is closed, and that its complement is the largest open nullset.

(2.9.4) Let $X$ be a compact Hausdorff space, and $\mu$ a Borel measure with $\mu(X)=1$.

(i) Show that $\mu(\operatorname{supp}\mu)=1$.

(ii) If $H\subsetneq\operatorname{supp}\mu$ is compact, show that $\mu(H)<1$.

(2.9.5) On $X=\mathbb{R}^2$, define

$$d((x_1,y_1),(x_2,y_2))=\begin{cases}|y_1-y_2|, & x_1=x_2\\ 1+|y_1-y_2|, & x_1\ne x_2\end{cases}$$

(i) Show that $d$ is a metric.

(ii) Show that $(X,d)$ is locally compact.

(iii) For $f \in C_c(X)$, show that there are only finitely many $x_1, \ldots, x_n$ such that $f(x_j, y) \neq 0$ for at least one $y$.

(iv) Let
$$\Lambda f = \sum_{j=1}^{n} \int_{-\infty}^{\infty} f(x_j, y)\, dy,$$
and show that $\Lambda : C_c(X) \to \mathbb{C}$ is linear and positive.

(v) Let $\mu$ be the measure corresponding to $\Lambda$ via Riesz–Markov. Let $E = \{(x, 0) : \ x \in \mathbb{R}\}$. Show that $\mu(E) = \infty$ and that $\mu(K) = 0$ for all $K \subset E$ compact. Does this contradict the Riesz–Markov theorem?

(2.9.6) Consider the same topological space $X$ from Exercise 2.9.5.

(i) Show that if $E \in \mathcal{B}(X)$, then each vertical slice $E_x$ is Borel.

(ii) Define
$$\mu(E) = \sum_{x} m(E_x), \qquad\qquad E \in \mathcal{B}(X).$$
Show that $\mu$ is a measure, inner regular.

(iii) Show that the $\mu$-nullsets are those Borel sets that intersect each vertical line in a nullset. Show that the diagonal $D = \{(x, x) : x \in \mathbb{R}\}$ is a nullset, and that every open $V$ with $D \subset V$ intersects each vertical line in a set of positive measure. Conclude that $\mu$ is not outer regular.

(iv) Define
$$\nu(E) = \begin{cases} \mu(E), & E \text{ intersects countably many vertical lines} \\ \infty, & \text{otherwise} \end{cases}$$
Show that $\nu$ is inner regular (on open sets), and outer regular on Borel sets.

(v) Show that
$$\int_X f\, d\mu = \int_X f\, d\nu, \qquad\qquad f \in C_c(X).$$

## 2.10. Complex Measures and Differentiation

As measures are supposed to "measure", we expect their values to be non-negative. But results like the Riesz–Markov Theorem emphasize the abstract point of view of seeing integrals as linear maps on the (compactly supported) continuous functions; it is then less obvious to expect them to be positive (in the sense of sending non-negative functions to non-negative numbers). For instance we have the linear map $f \longmapsto f(0) + f(1)$, say, which is positive; but the very similar map $f \longmapsto f(0) - f(1)$ is not positive. With this mindset, it makes sense to consider real-valued and even complex-valued measures, as these would correspond to such linear maps.

**2.10.1. Definition.**

*Let $(X, \mathcal{A})$ be a measurable space. A map $\mu : \mathcal{A} \to \mathbb{C}$ is a **complex measure** if*

$$\mu\Big(\bigcup_k E_k\Big) = \sum_k \mu(E_k) \tag{2.57}$$

*for every countable partition $\{E_k\}$ of $X$ by measurable sets.*

*If $\mu$ is as above but takes real values, we say that $\mu$ is a **signed measure**.*

*The **total variation** of $\mu$ is $|\mu| : \mathcal{A} \to [0, \infty]$ given by*

$$|\mu|(E) = \sup\Big\{ \sum_k |\mu(E_k)| : \{E_k\} \text{ measurable countable partition of } E \Big\}.$$

The equality (2.57) can only make sense if the convergence of the series is absolute, as the left-hand-side is immune to reordering (Exercise 2.10.1); this topic will be discussed in detail much later (cfr. Lemma 4.3.11 and the discussion preceding it, Remark 4.3.12, and subsection 9.8.4).

The notation $|\mu|$ suggests an absolute value, which we will see it somehow is. The first order of business, though, is to show that $|\mu|$ is a measure.

***2.10.2. Proposition.***

*The total variation $|\mu|$ is a measure.*

*Proof.* The condition (2.57) gives us $\mu(\varnothing) = \mu(\varnothing \cup \varnothing) = 2\mu(\varnothing)$, so $\mu(\varnothing) = 0$. Then $|\mu|(\varnothing) = 0$.

Note that the argument below works fine even if some of the terms are infinite. Fix a pairwise disjoint countable family $\{E_k\} \subset \mathcal{A}$. Write $E = \bigcup_k E_k$. Given numbers $t_k$ with $t_k < |\mu|(E_k)$, by definition there exist measurable partitions $\{F_{k,j}\}_j$ of each $E_k$ with $t_k < \sum_j |\mu(F_{k,j})|$. Using that $\{F_{k,j}\}_{k,j}$ is a measurable partition of $E$,

$$\sum_k t_k \leq \sum_k \sum_j |\mu(F_{k,j}| \leq |\mu|(E).$$

As this can be done for any choice of the $t_k$, this shows that $\sum\limits_k |\mu|(E_k) \leq |\mu|(E)$. Conversely, let $\{F_j\}$ be a measurable partition of $E$. Then, since $\{F_j \cap E_k\}$ is a partition for $E_k$,

$$\sum_j |\mu(F_j)| = \sum_j \Big| \sum_k \mu(F_j \cap E_k) \Big| \leq \sum_{j,k} |\mu(F_j \cap E_k)| = \sum_k |\mu|(E_k).$$

As this can be done for any partition $\{F_j\}$ of $E$, this shows that $|\mu|(E) \leq \sum\limits_k |\mu|(E_k)$, proving the equality. □

### *2.10.3. Lemma.*

*Let $z_1, \dots, z_n \in \mathbb{C}$. Then there exists $S \subset \{1, \dots, n\}$ such that*

$$\Big| \sum_{k \in S} z_k \Big| \geq \frac{1}{\pi} \sum_{k=1}^n |z_k|.$$

*Proof.* Write the polar form $z_k = |z_k|\, e^{i\gamma_k}$. Define sets $S(\theta) = \{k :\ \cos(\gamma_k - \theta) > 0\}$. These sets are for sure nonempty when $\theta$ is close enough to some $\gamma_k$, namely when $|\gamma_k - \theta| < \frac{\pi}{2}$. We have, for fixed $\theta$,

$$\begin{aligned} \Big| \sum_{k\in S(\theta)} z_k \Big| &= \Big| \sum_{k\in S(\theta)} e^{-i\gamma_k} |z_k| \Big| = \Big| \sum_{k\in S(\theta)} |z_k|\, e^{i(\gamma_k - \theta)} \Big| \\ &\geq \operatorname{Re} \sum_{k\in S(\theta)} |z_k|\, e^{i(\gamma_k-\theta)} = \sum_{k\in S(\theta)} |z_k|\, \cos(\gamma_k - \theta) \\ &= \sum_{k=1}^n |z_k|\, \cos^+(\gamma_k - \theta). \end{aligned}$$

The use of the positive part of the cosine allows us to sum over all indices $k$. As this last sum is a continuous expression on $\theta$, we can choose $\theta_0$ such that it is maximum. As such, it is greater than its average; so

$$\begin{aligned}
\Big|\sum_{k\in S(\theta_0)} z_k\Big| &\geq \frac{1}{2\pi}\int_0^{2\pi}\sum_{k=1}^n |z_k|\cos^+(\gamma_k-\theta)\,d\theta \\
&= \frac{1}{2\pi}\sum_{k=1}^n |z_k|\int_0^{2\pi}\cos^+(\gamma_k-\theta)\,d\theta \\
&= \frac{1}{2\pi}\sum_{k=1}^n |z_k|\int_0^{2\pi}\cos^+\theta\,d\theta \\
&= \frac{1}{2\pi}\sum_{k=1}^n |z_k|\int_{-\pi/2}^{\pi/2}\cos\theta\,d\theta \\
&= \frac{1}{\pi}\sum_{k=1}^n |z_k|. \qquad \square
\end{aligned}$$

The definition of complex measure gives us directly that $|\mu(E)| < \infty$ for all measurable $E$. What is interesting is that this implies that $|\mu|(X) < \infty$.

***2.10.4. Proposition.***

*If $\mu$ is a complex measure, then $|\mu|(X) < \infty$.*

*Proof.* Suppose that $E \in \mathcal{A}$ and $|\mu|(E) = \infty$. Put $t = \pi(1 + |\mu(E)|)$. Since $|\mu|(E) = \infty$, there exists a partition $\{E_j\}$ of $E$ and $N \in \mathbb{N}$ such that $t < \sum_{j=1}^N |\mu(E_j)|$. Applying Lemma 2.10.3 to the numbers $\mu(E_1), \dots, \mu(E_N)$, we obtain $S \subset \{1, \dots, N\}$. Define $A = \bigcup_{j\in S} E_j \subset E$. Then

$$|\mu(A)| = \Big|\sum_{j\in S}\mu(E_j)\Big| \geq \frac{1}{\pi}\sum_{j=1}^N |\mu(E_j)| > \frac{t}{\pi} > 1.$$

Also,

$$|\mu(E\setminus A)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| > \frac{t}{\pi} - |\mu(E)| = 1.$$

From $E = A \cup (E\setminus A)$ and writing $B = E\setminus A$, we obtained that $|\mu|(A) > 1$, $|\mu|(B) > 1$, and at least one of $|\mu|(A) = \infty$ or $|\mu|(B) = \infty$ holds (since $|\mu|(A\cup B) = \infty$).

If $|\mu|(X) = \infty$, by the above we can write $X = A_1 \cup B_1$, disjoint, with $|\mu(A_1)| > 1$ and $|\mu|(B_1) = \infty$. Now we apply the first paragraph to $B_1$, and we continue inductively. We obtain $A_1, A_2, \ldots$, pairwise disjoint, and with $|\mu(A_j)| > 1$ for all $j$. Then we would have

$$\mu\Big(\bigcup_j A_j\Big) = \sum_j \mu(A_j),$$

a contradiction since the sequence of numbers $\{\mu(A_j)\}$ satisfies $|\mu(A_j)| > 1$ for all $j$. So $|\mu|(X) < \infty$. □

It follows trivially from the definition that linear combinations of complex measures are again complex measures. That is, the set of complex measures on a fixed measurable space is a vector space. And $\|\mu\| = |\mu|(X)$ is a norm on that space.

The terminology we are using for complex measures causes that the usual measures from previous sections are not necessarily complex, if $\mu(X) = \infty$. So at times we will have to distinguish between complex measures and positive measures.

**2.10.5. Definition.**

*Let $(X, \mathcal{A})$ be a measurable space.*

*Given a signed measure $\mu : \mathcal{A} \to \mathbb{R}$, we define*

$$\mu^+ = \frac{1}{2}(\mu + |\mu|), \qquad \mu^- = \frac{1}{2}(|\mu| - \mu).$$

*Then $\mu = \mu^+ - \mu^-$ is the **Jordan Decomposition** of $\mu$.*

*If $\mu$ is a positive measure and $\lambda$ either a complex or a positive measure, we say that $\lambda$ is **absolutely continuous** with respect to $\mu$, denoted $\lambda \ll \mu$, if $\mu(E) = 0$ implies $\lambda(E) = 0$, $E \in \mathcal{A}$.*

*We say that a complex measure $\lambda$ is **concentrated** on $A$ if $\lambda(E) = \lambda(A \cap E)$ for all $E \in \mathcal{A}$. If there exist disjoint $A, B$ such that $\lambda_1$ and $\lambda_2$ are concentrated on $A, B$ respectively, we say that $\lambda_1$ and $\lambda_2$ are **mutually singular**, denoted $\lambda_1 \perp \lambda_2$.*

The measures $\mu^+$ and $\mu^-$ are mutually singular (Theorem 2.10.13). The Jordan decomposition is unique among decompositions $\mu = \mu_1 - \mu_2$ with $\mu_1, \mu_2$ positive and mutually singular (Corollary 2.10.14).

***2.10.6. Proposition.***

*With the above notation,*

*(i) if $\lambda$ is concentrated on $A$, then $|\lambda|$ is concentrated on $A$;*

*(ii) if $\lambda_1 \perp \lambda_2$, then $|\lambda_1| \perp |\lambda_2|$;*

*(iii) if $\lambda_1, \lambda_2 \perp \mu$, then $\lambda_1 + c\lambda_2 \perp \mu$ for all $c \in \mathbb{C}$;*

(iv) *if* $\lambda_1, \lambda_2 \ll \mu$, *then* $\lambda_1 + c\lambda_2 \ll \mu$ *for all* $c \in \mathbb{C}$;

(v) *if* $\lambda \ll \mu$, *then* $|\lambda| \ll \mu$;

(vi) *if* $\lambda_1 \ll \mu$ *and* $\lambda_2 \perp \mu$, *then* $\lambda_1 \perp \lambda_2$;

(vii) *if* $\lambda \ll \mu$ *and* $\lambda \perp \mu$, *then* $\lambda = 0$;

(viii) $|\lambda(E)| \leq |\lambda|(E)$ *for all* $E \in \mathcal{A}$. *In particular,* $\lambda \ll |\lambda|$.

*Proof.* Exercise 2.10.2. □

Let us see that the terminology *absolutely continuous* is justified.

***2.10.7. Proposition.***

*Let* $(X, \mathcal{A}, \mu)$ *be a measure space,* $\lambda$ *a complex measure on* $\mathcal{A}$. *The following statements are equivalent:*

(i) $\lambda \ll \mu$;

(ii) *for each* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that* $\mu(E) < \delta$ *implies* $|\lambda(E)| < \varepsilon$.

*Proof.* If (ii) holds and $\mu(E) = 0$, then $\mu(E) < \delta$ for any $\delta > 0$, and so $|\lambda(E)| < \varepsilon$ for an $\varepsilon > 0$. Thus $\lambda \ll \mu$.

Conversely, if (ii) fails, then there exists $\varepsilon > 0$ and $\{E_n\} \subset \mathcal{A}$ with $\mu(E_n) < 2^{-n}$ and $|\lambda(E_n)| \geq \varepsilon$. Put $A_m = \bigcup_{n\geq m} E_n$. Then $\mu(A_m) \leq \sum_{n\geq m} \mu(E_k) \leq \sum_{n\geq m} 2^{-n} = 2^{-m+1}$. Let $A = \bigcap_m A_m$. As $\{A_m\}$ is decreasing, we get by continuity of the measure that $\mu(A) = 0$. And

$$|\lambda|(A) = \lim_m |\lambda|(A_m) \geq \limsup_n |\lambda(E_n)| \geq \varepsilon.$$

So $|\lambda| \not\ll \mu$, and then $\lambda \not\ll \mu$. □

*2.10.8. Example.* Let $(X, \mathcal{A}, \mu)$ be a measure space and $f \in L^1(X)$. Then

$$\lambda(E) = \int_E f \, d\mu \tag{2.58}$$

defines a complex measure, and $\lambda \ll \mu$. We will see that, surprisingly, a converse to this situation holds.

For an example of a singular measure, consider the Lebesgue measure $m$ on $\mathbb{R}$. On $\mathcal{M}(\mathbb{R})$ define

$$\lambda(E) = \#(E \cap \mathcal{C}),$$

where $\mathcal{C}$ is the Cantor set and $\#$ denotes the counting measure. Then $\lambda$ is a positive measure and $\lambda \perp m$.

Forming a measure via (2.58) is so natural that physicists' intuition has historically treated every measure as such. That is the origin of the famous "Dirac function", which is in reality the Dirac measure—as as in Example 2.3.7—together with its associated linear functional as in (2.14). That is, physicists commonly write the "equality"

$$f(0) = \int_{\mathbb{R}} f(t)\,\delta(t)\,dt. \tag{2.59}$$

If such function $\delta$ existed so that (2.59) holds for every $f \in C_c(\mathbb{R})$, say, it would have to be "zero everywhere, infinite at 0, with $\int_{\mathbb{R}} \delta(t)\,dt = 1$", as the Dirac function is usually presented. But such a thing does not exist. The reason physicists do not go wrong when they think $d\delta = \delta(t)\,dt$ and they use (2.59) is that they never use the $\delta$ "function" as a function on its own. What they are doing is using the functional $f \longmapsto f(0)$ and (2.59) is just notation for the equation $f(0) = \displaystyle\int_{\mathbb{R}} f\,d\delta$. Their intuition is not totally unwarranted, though, as we will see soon.

**2.10.9. Definition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space and $\lambda$ a complex measure on $\mathcal{A}$. The* ***Lebesgue Decomposition*** *of $\lambda$ is*

$$\lambda = \lambda_a + \lambda_s,$$

*where $\lambda_a \ll \mu$ and $\lambda_s \perp \mu$.*

It is proper to call it "the" Lebesgue decomposition, as we will see right away. As an exception to the linear flow of the text, the proof below uses a result to be proven in a later chapter, the Riesz Representation Theorem (4.5.4). We consider Measure Theory as a pre-requisite for Functional Analysis, and so Chapter 4 had to be further along the book; but there is not logical inconsistency. The proof we use consists of a remarkable argument by von Neumann that joins the proof of the Lebesgue Decomposition and the existence of the Radon–Nikodym derivative. It is possible to prove both with just measure-theoretic arguments, and in that case avoid any forward reference; but we still prefer von Neumann's argument.

**2.10.10. Theorem. (Lebesgue–Radon–Nikodym)**

*Let $(X, \mathcal{A}, \mu)$ be a $\sigma$-finite measure space, and $\lambda$ a complex measure on $\mathcal{A}$. Then*

*(i) the Lebesgue Decomposition exists, and is unique. If $\lambda$ is positive and finite, so are $\lambda_a$ and $\lambda_s$;*

*(ii) there exists a unique $h \in L^1(\mu)$ such that $\lambda_a(E) = \int_E h\,d\mu$ for all $E \in \mathcal{A}$.*

*If $\lambda$ is instead a $\sigma$-finite measure (positive, possibly infinite) the result still holds, just with $h$ not necessarily integrable.*

*Proof.* Let us consider first the uniqueness in the Lebesgue Decomposition. If $\lambda_a^1+\lambda_s^1 = \lambda_a^2+\lambda_s^2$ with $\lambda_a^j \ll \mu$, $\lambda_s^j \perp \mu$, $j=1,2$, then $\lambda_a^1-\lambda_a^2 = \lambda_s^2-\lambda_s^1$. Using Proposition 2.10.6 we have $\lambda_a^1-\lambda_a^2 \ll \mu$, $\lambda_s^2-\lambda_s^1 \perp \mu$, so $\lambda_a^1-\lambda_a^2 = \lambda_s^2-\lambda_s^1 = 0$.

For the existence, assume first that $\lambda \geq 0$ and finite. Let $\omega \in L^1(X,\mu)$ be the function from Proposition 2.7.14. Define a positive measure

$$\eta(E) = \lambda(E) + \int_E \omega\,d\mu, \qquad E \in \mathcal{A}.$$

Note that $\eta$ is finite, since $\eta(X) = \lambda(X)+\|\omega\|_1$. Going first to simple functions and then to measurable nonnegative functions via Monotone Convergence, we have (note that we are using the same $\sigma$-algebra for all three measures)

$$\int_X f\,d\eta = \int_X f\,d\lambda + \int_X f\omega\,d\mu. \tag{2.60}$$

For any $f \in L^2(X,\eta)$, and since $\lambda \leq \eta$,

$$\Big|\int_X f\,d\lambda\Big| \leq \int_X |f|\,d\lambda \leq \int_X |f|\,d\eta \leq \|f\|_2\,\eta(X)^{1/2}.$$

This shows that the linear map $\varphi : L^2(X,\eta) \to \mathbb{C}$ given by $\varphi(f) = \int_X f\,d\lambda$ is bounded. By the Riesz Representation Theorem (4.5.4) there exists $g \in L^2(X,\eta)$ with

$$\int_X f\,d\lambda = \int_X fg\,d\eta, \qquad f \in L^2(X,\eta).$$

If $E \in \mathcal{A}$ and $\lambda(E) > 0$, then $1_E \in L^2(X,\eta)$ since $\eta$ is finite. Then

$$\lambda(E) = \int_X 1_E\,d\lambda = \int_X 1_E g\,d\eta. \tag{2.61}$$

Thus

$$0 \leq \frac{1}{\eta(E)}\int_E g\,d\eta = \frac{\lambda(E)}{\eta(E)} \leq 1.$$

As this can be done for any $E \in \mathcal{A}$ with $\eta(E) > 0$, we conclude that $0 \le g \le 1$ a.e.$(\eta)$; this was done in Exercise 2.5.22. Extending (2.61) to nonnegative measurable functions via simple functions and Monotone Convergence, for any positive measurable $f$ we have by (2.60)

$$\int_X f\,d\lambda = \int_X fg\,d\eta = \int_X fg\,d\lambda + \int_X fg\omega\,d\mu;$$

we may rewrite this as

$$\int_X (1-g)f\,d\lambda = \int_X fg\omega\,d\mu. \tag{2.62}$$

Let

$$A = \{0 \le g < 1\}, \qquad B = \{g = 1\},$$

and

$$\lambda_a(E) = \lambda(A \cap E), \qquad \lambda_s(E) = \lambda(B \cap E).$$

Putting $f = 1_B$ in (2.62) we get that $0 = \int_B \omega\,d\mu$; since $\omega > 0$ a.e., this implies that $\mu(B) = 0$. So $\lambda_s \perp \mu$. If now we put $f = (1 + g + \cdots + g^n)1_E$ in (2.62) we have

$$\int_E (1 - g^{n+1})\,d\lambda = \int_E g(1 + g + \cdots + g^n)\,\omega\,d\mu.$$

On $A$ we have that $g^n \to 0$ monotonically. So $1 - g^{n+1} \nearrow 1$. Using Monotone Convergence we obtain

$$\begin{aligned}\lambda_a(E) = \lambda(A \cap E) &= \lim_n \int_{A\cap E} (1 - g^{n+1})\,d\lambda = \lim_n \int_{A\cap E} \Big(\sum_{k=0}^{n} g^k\Big)\omega\,d\mu \\ &= \int_{A\cap E} \Big(\sum_{k=0}^{\infty} \omega g^n\Big)\,d\mu = \int_E h\,d\mu,\end{aligned}$$

with $h = 1_A \displaystyle\sum_{k=0}^{\infty} \omega g^k$. From $\lambda_a(X) = |\lambda(A)| < \infty$, we conclude that $h \in L^1(\mu)$. And from the equality $\lambda_a(E) = \int_E h\,d\mu$ it follows that $\lambda_a \ll \mu$.

When $\lambda$ is an arbitrary complex measure on $\mathcal{A}$, using the Jordan Decomposition we can write $\lambda = \lambda_1 - \lambda_2 + i(\lambda_3 - \lambda_4)$, with each of $\lambda_j$ positive and finite. By the above construction we obtain an $h_j$ for each $\lambda_j$, and then $h = (h_1 - h_2) + i(h_3 - h_4)$ is the function we need.

When $\lambda$ is a $\sigma$-finite infinite measure, writing $X = \displaystyle\bigcup_n X_n$ pairwise disjoint union with $\lambda(X_n) < \infty$ for all $n$ the whole proof applies verbatim on each $X_n$; then we form $h = \sum_n h_n 1_{X_n}$. ☐

When the measure $\mu$ is not $\sigma$-finite, there may not be a Radon–Nikodym derivative (Exercise 2.10.6). A trivial example is obtained by taking $X = \{1\}$, $\mu(X) = \infty$, $\lambda(X) = 1$. Then $\lambda \ll \mu$, but $\int_X h\, d\mu = \infty$ for any nonzero $h \geq 0$.

Theorem 2.10.10 is mostly applied in the case where $\lambda \ll \mu$, to obtain $h$. In such case this function $h$ is called the **Radon–Nikodym derivative** of $\lambda$ with respect to $\mu$, usually denoted by $\dfrac{d\lambda}{d\mu}$. In the above situation it is also common to write $d\lambda = h\, d\mu$.

Going back to the Dirac delta discussed above, we have $\delta \perp m$, where $\delta$ is the Dirac measure and $m$ the Lebesgue measure; thus the equality

$$\int_E f\, d\delta = \int_E f\, h\, dm$$

cannot hold for any measurable $h$. That is, writing $d\delta = \delta(t)\, dt$ is formally wrong, though it works for physicists for the reasons mentioned on page 217.

To understand the naming of the Radon–Nikodym derivative, consider the following example. Let $g : \mathbb{R} \to [0, \infty)$ be continuous and integrable. Define

$$\lambda(E) = \int_E g\, dm.$$

Using the Fundamental Theorem of Calculus, we have

$$g(x) = \frac{d}{dx} \lambda([a, x]).$$

The **polar decomposition** $z = re^{i\theta}$ of a complex number admits generalizations in several contexts, as we will see repeatedly. The theme is always that we decompose an object into an object of absolute value one times a positive object. Our first example of this comes next.

***2.10.11. Proposition. (Polar Decomposition)***

*Let $(X, \mathcal{A})$ be a measurable space and $\lambda$ a complex measure on $\mathcal{A}$. Then there exists a unique $h : X \to \mathbb{C}$, measurable, with $|h| = 1$ and such that $d\lambda = h\, d|\lambda|$.*

*Proof.* Since $\lambda \ll |\lambda|$, by Theorem 2.10.10 there exists a unique $h \in L^1(|\lambda|)$ with $d\lambda = h\, d|\lambda|$. Define, for $r > 0$, $A_r = \{|h| < r\}$. For a partition $\{E_k\}$ of $A_r$,

$$\sum_k |\lambda(E_k)| = \sum_k \left| \int_{E_k} h\, d|\lambda| \right| \leq \sum_k \int_{E_k} |h|\, d|\lambda| \leq r \sum_k |\lambda|(E_k) = r\, |\lambda|(A_r).$$

As this can be done for any partition of $A_r$, we obtain that $|\lambda|(A_r) \leq r\, |\lambda|(A_r)$. So, if $r < 1$, we have $|\lambda|(A_r) = 0$. Thus $|h| \geq 1$ a.e. ($|\lambda|$).

Conversely, if $|\lambda|(E) > 0$, then

$$\left| \frac{1}{|\lambda|(E)} \int_E h \, d|\lambda| \right| = \frac{|\lambda(E)|}{|\lambda|(E)} \leq 1.$$

Thus $|h| \leq 1$ a.e. ($|\lambda|$) by Exercise 2.5.22, proving that $|h| = 1$ a.e. □

For a first application of the Polar Decomposition, we have

**2.10.12. Proposition.**

*Let $(X, \mathcal{A}, \mu)$ be a measure space, $g \in L^1(\mu)$, and $\lambda$ the complex measure $\lambda(E) = \int_E g \, d\mu$. Then*

$$|\lambda|(E) = \int_E |g| \, d\mu, \qquad E \in \mathcal{A}.$$

*Proof.* By Proposition 2.10.11 there exists $h$ measurable, with $|h| = 1$ a.e. ($|\lambda|$) and such that $d\lambda = h \, d|\lambda|$. Then, for any $E \in \mathcal{A}$,

$$\int_X 1_E \, g \, d\mu = \int_E g \, d\mu = \lambda(E) = \int_E 1 \, d\lambda = \int_E h \, d|\lambda| = \int_X 1_E \, h \, d|\lambda|.$$

By linearity the above equality holds for all simple functions, and by Monotone Convergence and linearity for all measurable functions. We can thus apply the equality to $1_E \, \bar{h}$ to get

$$\int_E \bar{h} \, g \, d\mu = \int_E \bar{h} \, h \, d|\lambda| = \int_E |h|^2 \, d|\lambda| = |\lambda|(E)$$

From the fact that $\mu$ and $|\lambda|$ are nonnegative, $\bar{h}g \geq 0$ a.e. ($\mu$). So $\bar{h} \, g = |\bar{h} \, g| = |g|$. □

We already had a way to write a real finite measure as a difference of two positive measures in the Jordan decomposition. The Hahn decomposition below allows us to decompose the measure space in a compatible way.

**2.10.13. Theorem. (Hahn Decomposition)**

*Let $(X, \mathcal{A})$ be a measurable space and $\mu$ a signed measure on $\mathcal{A}$. Then there exist $A, B \in \mathcal{A}$ with $A \cup B = X$, $A \cap B = \varnothing$, and such that*

$$\mu^+(E) = \mu(A \cap E), \qquad \mu^-(E) = -\mu(B \cap E), \qquad E \in \mathcal{A}.$$

*The decomposition is unique in the sense that if $A', B'$ satisfy the same conditions, then $A \setminus A'$, $A' \setminus A$, $B \setminus B'$, and $B' \setminus B$ are nullsets for $|\mu|$.*

*Proof.* By the Polar Decomposition there exists $h$ measurable with $d\mu = h\,d|\mu|$, and $|h| = 1$ a.e. ($|\mu|$). As $\mu$ is real, so is $h$ almost everywhere by Exercise 2.5.22. By redefining $h$ on a nullset we may assume that $h$ is real everywhere; so $h$ only takes the values 1 and $-1$. Let $A = \{h = 1\}$ and $B = \{h = -1\}$. Then

$$\mu^+(E) = \frac{1}{2}\,(|\mu| + \mu)(E) = \frac{1}{2}\int_E (1+h)\,d|\mu| = \int_{E\cap A} h\,d|\mu| = \mu(A\cap E).$$

And then

$$\mu^-(E) = (\mu^+ - \mu)(E) = \mu(A\cap E) - \mu(E) = -\mu(E\setminus A\cap E) = -\mu(B\cap E).$$

For the uniqueness, suppose that $A', B' \in \mathcal{A}$ satisfy that $A'\cup B' = X$, $A'\cap B' = \varnothing$, and $\mu(A'\cap E) \geq 0$, $\mu(B'\cap E) \leq 0$ for all $E \in \mathcal{A}$. We have

$$0 \leq \mu(A\cap B') = \mu(B'\cap A) \leq 0.$$

Thus $\mu(A\setminus A') = \mu(A\cap B') = 0$. And $\mu^+(A\setminus A') = \mu(A\setminus A') = 0$. So $|\mu|(A\setminus A') = (2\mu^+ - \mu)(A\setminus A') = 0$. A similar reasoning works for the other three set differences. $\square$

With the Hahn Decomposition available, we can show the uniqueness of the Jordan Decomposition.

**2.10.14. Corollary.** *(Jordan Decomposition of a Measure)*

*Let $(X, \mathcal{A})$ be a measurable space and $\mu$ a signed measure on $\mathcal{A}$. The measures $\mu^+$ and $\mu^-$ are mutually singular. The Jordan Decomposition is unique, in the sense that if $\mu = \mu_1 - \mu_2$ with $\mu_1, \mu_2$ positive and mutually singular, then $\mu_1 = \mu^+$ and $\mu_2 = \mu^-$.*

*Proof.* Let $A, B$ be a Hahn decomposition of $X$. By Theorem 2.10.13 the measure $\mu^+$ is supported on $A$ and the measure $\mu^-$ is supported on $B$, so $\mu^+ \perp \mu^-$.

Now suppose that $\mu = \mu_1 - \mu_2$ with $\mu_1, \mu_2$ positive and mutually singular. Then there exist $A', B' \in \mathcal{A}$ such that $\mu_1(E) = \mu_1(A'\cap E)$, $\mu_2(E) = \mu_2(B'\cap E)$ for all $E \in \mathcal{A}$. By the uniqueness of the Hahn decomposition, the set differences $A\setminus A'$, $A'\setminus A$, $B\setminus B'$, and $B'\setminus B$ are nullsets for $|\mu|$.

We have, for $E \in \mathcal{A}$, and $N \in \mathcal{A}$ a nullset for $|\mu|$,

$$|\mu(N\cap E)| \leq |\mu|(N\cap E) \leq |\mu|(N) = 0.$$

Then

$$\begin{aligned}\mu_1(E) &= \mu_1(A'\cap E) = \mu_1(A'\cap E) - \mu_2(A'\cap E) = \mu(A'\cap E)\\ &= \mu\big([(A'\setminus A)\cup(A'\cap A)]\cap E\big) = \mu\big((A'\setminus A)\cap E\big) + \mu\big((A'\cap A)\cap E\big)\\ &= \mu\big((A'\cap A)\cap E\big) = \mu\big((A'\cap A)\cap E\big) + \mu\big((A\setminus A')\cap E\big)\end{aligned}$$

$$= \mu(A \cap E) = \mu^+(E).$$

And then $\mu_2(E) = \mu_1(E) - \mu(E) = \mu^+(E) - \mu(E) = \mu^-(E)$. □

*2.10.15. Remark*. For a signed measure we have $|\mu| = \mu^+ + \mu^-$. There is no similar formula for the total variation of complex measures, though. Write $\mu = \mu_r + i\mu_i$ with $\mu_r, \mu_i$ real. By Proposition 2.10.11, we can write $\mu(E) = \int_X h\, d|\mu|$ for all $E \in \mathcal{A}$ and $|h| = 1$. Then

$$\mu_r(E) = \operatorname{Re} \mu(E) = \int_E \operatorname{Re} h\, d|\mu|, \qquad \mu_i(E) = \operatorname{Im} \mu(E) = \int_E \operatorname{Im} h\, d|\mu|.$$

Using Proposition 2.10.12,

$$|\mu_r|(E) = \int_E |\operatorname{Re} h|\, d|\mu|, \qquad |\mu_i|(E) = \int_E |\operatorname{Im} h|\, d|\mu|.$$

So we cannot expect some formula (like addition in the real case) on $|\mu_r|$ and $|\mu_i|$ to give us $|\mu|$, as we are free to choose $\operatorname{Re} h$ and $\operatorname{Im} h$, as long as their squares add to 1. The difference with the real case is that for $h$ real we have $|h| = h^+ + h^-$ and the dependence is linear, which is preserved by the integrals. But when $h$ is complex $|h|^2 = |\operatorname{Re} h|^2 + |\operatorname{Im} h|^2$ and linearity is lost.

At this stage, we have still not defined what the integral against a complex measure is. Without much imagination,

**2.10.16. Definition. (Lebesgue Integral against a Complex Measure)**

*Let $(X, \mathcal{A})$ be a measurable space and $\lambda$ a complex measure on $\mathcal{A}$. Let $g : X \to \mathbb{C}$ be measurable. Using the Jordan decomposition, let $\lambda = \lambda_1 - \lambda_2 + i(\lambda_3 - \lambda_4)$ be the unique decomposition where $\lambda_j$ is a positive measure for $j = 1, 2, 3, 4$, and $\lambda_1 \perp \lambda_2$, $\lambda_3 \perp \lambda_4$. Then we define*

$$\int_X g\, d\lambda = \int_X g\, d\lambda_1 - \int_X g\, d\lambda_2 + i\left(\int_X g\, d\lambda_3 - \int_X g\, d\lambda_4\right),$$

*if the four integrals on the right exist.*

Besides linearity—which follows from a straightforward computation—the crucial property we need from integrals is the triangle inequality. In this case,

***2.10.17. Proposition.***

*Let $(X, \mathcal{A})$ be a measurable space, $\lambda$ a complex measure on $\mathcal{A}$, and $g : X \to \mathbb{C}$ measurable. Then*

$$\left|\int_X g\,d\lambda\right| \leq \int_X |g|\,d|\lambda|,$$

*provided that the integral on the left exists.*

*Proof.* Let $c \in \mathbb{C}$ such that $c\int_X g\,d\lambda = \left|\int_X g\,d\lambda\right|$. Then $|c| = 1$. By the Polar Decomposition (Proposition 2.10.11) and Exercise 2.10.9 there exists measurable $h$ with $|h| = 1$ and such that $\int_X f\,d\lambda = \int_X fh\,d|\lambda|$. Then

$$\left|\int_X g\,d\lambda\right| = \int_X cg\,d\lambda = \int_X cgh\,d|\lambda| \leq \int_X |cgh|\,d|\lambda| = \int_X |g|\,d|\lambda|. \qquad \square$$

**2.10.1. Exercises.**

(2.10.1) Let $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ such that $\sum_n c_n$ converges. Show that $\sum_n c_n$ converges absolutely if and only if the limit does not change under permutations; concretely, show that the following statements are equivalent:

(i) $\sum_n c_n$ converges absolutely;

(ii) $\sum_n c_{\sigma(n)} = \sum_n c_n$ for all permutations $\sigma : \mathbb{N} \to \mathbb{N}$;

(iii) $\sum_n c_{\sigma(n)}$ converges for all permutations $\sigma : \mathbb{N} \to \mathbb{N}$.

(2.10.2) Prove Proposition 2.10.6.

(2.10.3) Let $\mu$ be a complex measure on a $\sigma$-algebra $\mathcal{A}$. For $E \in \mathcal{A}$, define

$$\lambda(E) = \sup\left\{\sum_{j=1}^m |\mu(E_j)| : \ E_j \in \mathcal{A} \ \text{ disjoint}, \ \bigcup_j E_j = E\right\}.$$

Show that $\lambda = |\mu|$.

(2.10.4) Let $\lambda_1, \lambda_2$ be mutually singular complex measures on a $\sigma$-algebra $\mathcal{A}$ over $X$. Show that

$$|\lambda_1 + \lambda_2| = |\lambda_1| + |\lambda_2|.$$

(2.10.5) Use Proposition 2.10.7 for an alternative proof of Exercise 2.5.8.

(2.10.6) *(Radon–Nikodym and Lebesgue's Decomposition can fail when $X$ is not $\sigma$-finite)* Let $\mu$ be the Lebesgue measure on $(0, 1)$ and $\lambda$ the counting measure on $\mathcal{M}((0, 1))$. Show that $\lambda$ has no Lebesgue decomposition relative to $\mu$, and that $\mu \ll \lambda$ and $\mu$ is bounded, but there is no measurable function $h$ such that $d\mu = h\,d\lambda$.

(2.10.7) Suppose that $\{g_n\}$ is a sequence of positive continuous functions on $[0,1]$. Let $\mu$ be a positive Borel measure on $[0,1]$, and suppose also that

(i) $\lim_{n\to\infty} g_n(x) = 0$ a.e.[m]

(ii) $\displaystyle\int_{[0,1]} g_n\,dm = 1$ for all $n$

(iii) $\displaystyle\lim_{n\to\infty}\int_{[0,1]} fg_n\,dm = \int_{[0,1]} f\,d\mu$ for all $f \in C[0,1]$.

Does it follow that $\mu \perp m$?

*(Hint: Egorov's Theorem. There are probably other ways to attack this)*

(2.10.8) Let $(X,\mathcal{A})$ be a measurable space, and $\mu,\lambda$ positive measures such that $\mu$ is $\sigma$-finite and $\lambda \ll \mu$. Let $g = d\lambda/d\mu$. Show that for any measurable $f : X \to \mathbb{C}$ such that $\int_X f\,d\lambda$ exists,

$$\int_X f\,d\lambda = \int_X fg\,d\mu. \tag{2.63}$$

(2.10.9) Let $(X,\mathcal{A},\mu)$ be a finite measure space and $f \in L^1(X)$. Define $\nu : \mathcal{A} \to \mathbb{C}$ by

$$\nu(E) = \int_E f\,d\mu.$$

(i) Show that $\nu$ is a complex measure on $(X,\mathcal{A})$ with $\nu \ll \mu$.

(ii) Show that for any measurable function $g : X \to \mathbb{C}$,

$$\int_X g\,d\nu = \int_X gf\,d\mu \tag{2.64}$$

if the left integral exists.

(iii) Conclude that $f = d\nu/d\mu$ a.e. $(\mu)$.

(2.10.10) Let $(X,\mathcal{A})$ be a measurable space, and let $\mu,\lambda$ be $\sigma$-finite positive measures such that $\lambda \ll \mu$ and $\mu \ll \lambda$. Show that

$$\frac{d\lambda}{d\mu} = \left(\frac{d\mu}{d\lambda}\right)^{-1} \qquad \text{a.e. } (\mu,\lambda).$$

(2.10.11) If $d\nu = f\,d\mu$ and $d\lambda = g\,d\nu$ for positive measures $\mu$ and $\nu$, show that $\lambda \ll \mu$ and find the Radon–Nikodym derivative $d\lambda/d\mu$.

(2.10.12) Let $(X,\Sigma)$ be a measurable space and $M(X)$ the set of complex measures on $\Sigma$. Show that, with the norm $\|\mu\| = |\mu|(X)$ and the obvious addition and multiplication by scalars, $M(X)$ is complete with the metric $d(\mu,\eta) = \|\mu - \eta\|$.

(2.10.13) Let $(X,\mathcal{A},\mu)$ be a $\sigma$-finite measure space, and $\lambda_1,\lambda_2$ positive $\sigma$-finite measures with $\lambda_1 \ll \mu$ and $\lambda_2 \ll \mu$. Give a necessary and sufficient condition, in terms of the Radon–Nikodym derivatives, for $\lambda_1 \ll \lambda_2$. In such case, express $d\lambda_1/d\lambda_2$ in terms of $d\lambda_1/d\mu$ and $d\lambda_2/d\mu$.

(2.10.14) Let $(X, \Sigma, \mu)$ be a $\sigma$-finite measure space. Show that there exists a finite measure $\nu$ on $\Sigma$ such that $\nu \ll \mu$ and $\mu \ll \nu$. Does such a $\nu$ always exist when $\mu$ is not necessarily $\sigma$-finite?

(2.10.15) Show that a linear combination of absolutely continuous functions is again absolutely continuous.

(2.10.16) Assume that both $f$ and $Mf$ are in $L^1(\mathbb{R}^n)$. Prove that $f = 0$ a.e.

(2.10.17) Let $X$ be a locally compact Hausdorff space. We say that a complex measure $\mu$ on $\mathcal{B}(X)$ is **regular** if $|\mu|$ is regular, and **Radon** if $|\mu|$ is Radon. Let $\gamma$ be a positive Radon measure on $\mathcal{B}(X)$, $g \in L^\infty(X)$, and define a complex Borel measure $\mu$ by

$$\mu(E) = \int_E g\, d\gamma.$$

Show that $\mu$ is regular if $\gamma$ is regular, and Radon if $\gamma$ is Radon.

(2.10.18) Let $X \subset \mathbb{C}$ be compact and $\mu$ a complex Radon measure on $X$. Show that if $\int_X f\, d\mu \geq 0$ for all polynomials $f$ such that $f(X) \subset [0, \infty)$, then $\mu$ is a positive measure.

(2.10.19) Let $X$ be a compact Hausdorff space and $\mu$ a regular complex Borel measure on $\mathcal{A}$. Let $\varphi : C(X) \to \mathbb{C}$ be given by $\varphi(f) = \int_X f\, d\mu$. Show that $\varphi$ is multiplicative if and only if $\mu = \delta_{x_0}$ for some $x_0 \in X$.

## 2.11. Differentiation

We now turn to differentiation. As opposed to the previous section, this section is not very relevant for the rest of the book, but it shows some of the connections from measure theory with earlier ideas from calculus.

We will restrict ourselves to $\mathbb{R}^n$. The theory can be extended beyond, to some metric spaces, but probably not much more. The reason being that one needs to be able to put some constraints on how sets shrink, as we will see soon.

**2.11.1. Definition.**

*Let $\mu$ be a complex Borel measure on $\mathbb{R}^n$. The **symmetric derivative** of $\mu$ at $x \in \mathbb{R}^n$ is (if it exists)*

$$(D\mu)x = \lim_{r \to 0} \frac{\mu(B_r(x))}{m(B_r(x))}.$$

The ***maximal function*** *for $\mu$ at $x$ is*

$$(M\mu)x = \sup\left\{\frac{|\mu|(B_r(x))}{m(B_r(x))} \,:\, r > 0\right\}.$$

The name "derivative" for $D\mu$ can be understood in light of the following:

***2.11.2. Proposition.***

*Let $\mu$ be a complex Borel measure on $\mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{C}$ be $f(t) = \mu(-\infty, t)$. The following statements are equivalent:*

*(i) $f$ is differentiable at $t$ and $f'(t) = c$;*

*(ii) for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $I$ is an interval with $t \in I$ and $m(I) < \delta$, then*

$$\left|\frac{\mu(I)}{m(I)} - c\right| < \varepsilon.$$

*Proof.* If $I = [t - \delta_1, t + \delta_2]$ with $\delta_1, \delta_2 > 0$, then

$$\begin{aligned}\frac{\mu(I)}{m(I)} &= \frac{f(t+\delta_2) - f(t-\delta_1)}{\delta_1 + \delta_2} \\ &= \frac{f(t+\delta_2) - f(t)}{\delta_2}\,\frac{\delta_2}{\delta_1+\delta_2} + \frac{f(t-\delta_1) - f(t)}{-\delta_1}\,\frac{\delta_1}{\delta_1+\delta_2}\end{aligned}$$

As this last term is a convex combination of Newton quotients of $f$ at $t$, this shows that if $f$ is differentiable at $t$ then $\mu(I)/m(I) \to f'(t)$ as $m(I) \to 0$. For the converse, if we take $I = [t, t+h)$ then $\mu(I)/m(I)$ is precisely the Newton quotient for $f$ at $t$. □

***2.11.3. Proposition.***

*$M\mu$ is lower semicontinuous. In particular, it is measurable.*

*Proof.* Let $E = (M\mu)^{-1}(\alpha, \infty)$. We may assume $\alpha > 0$, as $(M\mu)x \geq 0$. We will show that $E$ is open. Fix $x \in E$; as $(M\mu)x > \alpha$, there exists $r > 0$ such that $|\mu|(B_r(x)) = t\, m(B_r(x))$ for some $t > \alpha$. Since $t/\alpha > 1$, there exists $\delta > 0$ with $(r+\delta)^n < r^n\, t/\alpha$. Let $y \in B_\delta(x)$. As $B_r(x) \subset B_{r+\delta}(y)$,

$$\begin{aligned}|\mu|(B_{r+\delta}(y)) &\geq |\mu|(B_r(x)) = t\, m(B_r(x)) = t\, m(B_r(y)) \\ &= t\left(\frac{r}{r+\delta}\right)^n m(B_{r+\delta}(y)) > \alpha m(B_{r+\delta}(y)).\end{aligned}$$

Thus $(M\mu)y > \alpha$, showing that $B_\delta(x) \subset E$ and so $E$ is open. □

The next lemma will be crucial, and it shows the kind of manipulation that we would not be able to do on an arbitrary measure space, or even an arbitrary locally compact Hausdorff space.

### *2.11.4. Lemma.*

*Let $X$ be a metric space. If $W = \bigcup_{j=1}^{N} B_{r_j}(x_j)$, then there exists $S \subset \{1, \ldots, N\}$ such that $\{B_{r_j}(x_j) : j \in S\}$ are pairwise disjoint and $W \subset \bigcup_{j \in S} B_{3r_j}(x_j)$.*

*Proof.* By reordering if necessary, we assume that $r_1 \geq \cdots \geq r_N$. Let $j_1 = 1$. Now we discard each $j$ such that $B_{r_j}(x_j) \cap B_{r_1}(x_1) \neq \varnothing$. Let $j_2$ the largest index remaining. Now repeat, until we end up with $S = \{j_1, \ldots, j_\ell\}$. These remaining balls are pairwise disjoint by construction.

Every discarded $B_{r_j}(x_j)$ satisfies $B_{r_j}(x_j) \cap B_{r_h}(x_h) \neq \varnothing$ for some $h \in S$ with $r_j \leq r_h$. For any $x \in B_{r_j}(x_j)$, choose $y \in B_{r_j}(x_j) \cap B_{r_h}(x_h)$. Then

$$d(x, r_h) \leq d(x, x_j) + d(x_j, y) + d(y, r_h) \leq r_j + r_j + r_h \leq 3r_h.$$

Thus $B_{r_j}(x_j) \subset B_{3r_h}(x_h)$. □

Recall that if $\mu$ is a complex measure on $\mathbb{R}^n$ then $|\mu|(\mathbb{R}^n) < \infty$ by Proposition 2.10.4.

### *2.11.5. Proposition.*

*Let $\mu$ be a complex measure on $\mathbb{R}^n$, $\alpha > 0$. Then*

$$m(\{M\mu > \alpha\}) \leq 3^n \, \alpha^{-1} \, |\mu|(\mathbb{R}^n).$$

*Proof.* We proved in Proposition 2.11.3 that $\{M\mu > \alpha\}$ is open. Let $K \subset \{M\mu > \alpha\}$ be compact. For each $x \in K$ we have $M\mu(x) > \alpha$. By definition there exists $r_x > 0$ such that $|\mu|(B_{r_x}(x)) > \alpha \, m(B_{r_x}(x))$. By compactness there exists $x_1, \ldots, x_N \in K$ such that $K \subset \bigcup_{j=1}^{N} B_{r_{x_j}}(x_j)$. Using Lemma 2.11.4 we get $S \subset \{1, \ldots, N\}$ such that $K \subset \bigcup_{j \in S} B_{3r_{x_j}}(x_j)$ and the balls $B_{r_{x_j}}$ are pairwise disjoint. Then

$$m(K) \leq \sum_{j \in S} m(B_{3r_{x_j}}(x_j)) = 3^n \sum_{j \in S} m(B_{r_{x_j}}(x_j))$$

$$\leq \frac{3^n}{\alpha} \sum_{j \in S} |\mu|(B_{r_{x_j}}(x_j)) \leq \frac{3^n}{\alpha} |\mu|(\mathbb{R}^n),$$

where we use in the last inequality that the balls corresponding to the indices in $S$ are pairwise disjoint. Now

$$m(\{M\mu > \alpha\}) = \sup\{m(K) : \; K \subset \{M\mu > \alpha\}, \;\; \text{compact}\} \leq 3^n \alpha^{-1} |\mu|(\mathbb{R}^n).$$

□.

**2.11.6. Definition.**

*Let $f : \mathbb{R}^n \to \mathbb{C}$ be Borel measurable. We say that $f \in L^1_w(\mathbb{R}^n)$ ("$L^1$-weak") if the function $\alpha \longmapsto \alpha\, m(\{|f| > \alpha\})$, $\alpha > 0$, is bounded.*

If $f \in L^1(\mathbb{R}^n)$ and $\alpha > 0$, then

$$\alpha\, m(\{|f| > \alpha\}) = \int_{m(\{|f|>\alpha\})} \alpha\, dm \leq \int_{\mathbb{R}^n} |f|\, dm = \|f\|_1.$$

So $L^1(\mathbb{R}^n) \subset L^1_w(\mathbb{R}^n)$. But the inclusion is proper. For instance consider, when $n = 1$, the function $f(t) = \frac{1}{t} 1_{(0,1)}$. Then $f \notin L^1(\mathbb{R})$, but

$$m(\{|f| > \alpha\}) = m(\{\frac{1}{t} > \alpha\}) = m\big((0, \frac{1}{\alpha})\big) = \frac{1}{\alpha},$$

so $f \in L^1_w(\mathbb{R})$.

**2.11.7. Definition.**

*For $f \in L^1(\mathbb{R}^n)$, its **maximal function** $Mf : \mathbb{R}^n \to [0, \infty]$ is*

$$(Mf)(x) = \sup_{r>0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f|\, dm.$$

*An $x \in \mathbb{R}^n$ is a **Lebesgue point** for $f$ if*

$$\lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(t) - f(x)|\, dm(t) = 0.$$

*2.11.8. Remark.* When $d\mu = f\, dm$, we have the equality $M\mu = Mf$. Then we see, by Proposition 2.11.5 that the map $f \longmapsto Mf$ maps $L^1(\mathbb{R}^n)$ to $L^1_w(\mathbb{R}^n)$.

Regarding Lebesgue points, if $x$ is a Lebesgue point for $f$ we have

$$f(x) = \lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} f\, dm. \tag{2.65}$$

When $f$ is continuous at $x$, it is trivial to check that $x$ is a Lebesgue point for $f$.

**2.11.9. Theorem. (Lebesgue Points)**

*Let $f \in L^1(\mathbb{R}^n)$. Then almost every $x \in \mathbb{R}^n$ is a Lebesgue point for $f$.*

*Proof.* Define

$$g_r(x) = \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(t) - f(x)|\, dm(t), \qquad r > 0,$$

and

$$g(x) = \limsup_{r\to 0} g_r(x).$$

At every point where $g(x) = 0$ we have that $x$ is a Lebesgue point for $f$. So the goal is to show that $g = 0$ a.e.

Fix $k \in \mathbb{N}$. By Proposition 2.8.18 there exists $f_k \in C_c(\mathbb{R}^n)$ with $\|f - f_k\|_1 < \frac{1}{k}$. Then

$$\begin{aligned} g_r(x) &= \frac{1}{m(B_r(x))} \int_{B_r(x)} |f(t) - f(x)|\, dm(t) \\ &\le \frac{1}{m(B_r(x))} \int_{B_r(x)} |(f - f_k)(t) - (f - f_k)(x)|\, dm(t) \\ &\qquad + \frac{1}{m(B_r(x))} \int_{B_r(x)} |f_k(t) - f_k(x)|\, dm(t). \end{aligned}$$

Since $f_k$ is continuous, every point is a Lebesgue point; so

$$\begin{aligned} g(x) &\le \frac{1}{m(B_r(x))} \int_{B_r(x)} |(f - f_k)(t) - (f - f_k)(x)|\, dm(t) \\ &\le (M(f - f_k))(x) + |f - f_k|(x). \end{aligned}$$

If $\beta > 0$ and $g(x) > 2\beta$, then at least one of the two terms above has to be greater than $\beta$. Thus, for any $k$,

$$\{g > 2\beta\} \subset \{M(f - f_k) > \beta\} \cup \{|f - f_k| > \beta\}.$$

Using Proposition 2.11.5 and the fact that $M\mu = M(f - f_k)$ when $d\mu = (f - f_k)\, dm$,

$$\begin{aligned} m(\{M(f - f_k) > \beta\}) + m(\{|f - f_k| > \beta\}) & \\ &\le 3^n \beta^{-1} \|f - f_k\|_1 + \beta^{-1} \|f - f_k\|_1 \\ &\le \frac{3^n + 1}{k\beta}. \end{aligned}$$

As
$$\{g > 2\beta\} \subset \bigcap_k \left[\{M(f - f_k) > \beta\} \cup \{|f - f_k| > \beta\}\right]$$
and the right-hand-side is a nullset by continuity of the measure, we conclude that $\{g > 2\beta\}$ is measurable and a nullset by completeness. Thus $g = 0$ a.e. □

Let us now obtain some consequences of Theorem 2.11.9. The first one we promised on Remark 2.8.23.

**2.11.10. Corollary.**

*Let $f \in L^p(\mathbb{R}^d)$, $g \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ with $\|g\|_1 = 1$. Then, with the notation of Proposition 2.8.22, $f * g_n(x) \to f(x)$ almost everywhere.*

*Proof.* Let $c > 0$ such that $m(B_r(0)) = cr^d$. From the estimate in the proof of (iii) in Proposition 2.8.22 we have
$$\begin{aligned}
|f * g_n(x) - f(x)| &\leq \int_{\mathbb{R}^d} |f(x-t) - f(x)|^p \, |g_n(t)| \, dt \\
&= \frac{1}{(1/n)^d} \int_{\mathbb{R}^n} |f(x-t) - f(x)|^p \, |g(nt)| \, dt \\
&\leq \frac{\|g\|_\infty}{(1/n)^d} \int_{\mathbb{R}^n} |f(x-t) - f(x)|^p \, dt \\
&= \frac{c\|g\|_\infty}{m(B_{1/n}(0))} \int_{\mathbb{R}^n} |f(x-t) - f(x)|^p \, dt.
\end{aligned}$$
By Theorem 2.11.9, the expression on the right converges to zero almost everywhere. □

**2.11.11. Corollary.**

*Let $\mu$ be a complex Borel measure on $\mathbb{R}^n$ such that $\mu \ll m$. Then $D\mu = \frac{d\mu}{dm}$ a.e. $(m)$ and*
$$\mu(E) = \int_E (D\mu) \, dm, \qquad \text{for all } E \in \mathcal{B}(\mathbb{R}^n).$$

*Proof.* If $x$ is a Lebesgue point for $\frac{d\mu}{dm}$, then
$$\frac{d\mu}{dm}(x) = \lim_{r \to 0} \frac{1}{m(B_r(x))} \int_{B_r(x)} \frac{d\mu}{dm} \, dm = \lim_{r \to 0} \frac{\mu(B_r(x))}{m(B_r(x))} = (D\mu)x.$$ □

We cannot expect to have an analog of Theorem 2.11.9 if we use arbitrary sets instead of balls. But the result does work if we use sets that "shrink nicely".

**2.11.12. Definition.**

*Let $x \in \mathbb{R}^n$. A sequence $\{E_k\}$ of Borel sets* ***shrinks nicely*** *to $x$ if there exists $\alpha > 0$ and a sequence $\{r_k\} \subset (0, \infty)$ with $r_k \to 0$, such that $E_k \subset B_{r_k}(x)$ and $m(E_k) \geq \alpha\, m(B_{r_k}(x))$.*

For instance, the sets

$$E_k = \left\{(x, y) : \ \frac{1}{2k} \leq x \leq \frac{1}{k},\ \frac{1}{25k} \leq y \leq \frac{1}{9k}\right\}$$

shrink nicely to 0 on $\mathbb{R}^2$; it is enough to take $r_k = 2/k$ and $\alpha = \frac{2}{225\pi}$. It is interesting that $0 \notin E_k$ even though the sets shrink to 0. Meanwhile, the sets

$$F_k = \left\{(x, y) : \ -\frac{1}{k} \leq x \leq \frac{1}{k},\ -\frac{1}{k^2} \leq y \leq \frac{1}{k^2}\right\}$$

do **not** shrink nicely to 0. To satisfy $F_k \subset B_{r_k}(0)$ we need $r_k \geq \frac{1}{k}$. But then $\alpha$ would have to satisfy

$$\frac{4}{k^3} = m(F_k) \geq \alpha\, m(B_{r_k}(0)) = \alpha\, \pi\, r_k^2 \geq \frac{\alpha\, \pi}{k^2}.$$

This is equivalent to $\alpha \leq \frac{4}{k\pi}$ for all $k$, which is impossible since we need $\alpha > 0$.

***2.11.13. Theorem. (Lebesgue Differentiation)***

*Let $f \in L^1(\mathbb{R}^n)$. If $x$ is a Lebesgue point for $f$ and $\{E_k\}$ shrinks nicely to $x$, then*

$$f(x) = \lim_{k\to\infty} \frac{1}{m(E_k)} \int_{E_k} f\, dm.$$

*Proof.* Since $E_k \subset B_{r_k}(x)$ and $m(E_k) \geq \alpha\, m(B_{r_k}(x))$,

$$0 \leq \frac{1}{m(E_k)} \int_{E_k} |f - f(x)|\, dm \leq \frac{1}{\alpha} \frac{1}{m(B_{r_k}(x))} \int_{B_{r_k}(x)} |f - f(x)|\, dm \xrightarrow[k\to\infty]{} 0.$$

□

To see how things can go wrong when sets do not shrink nicely, we mention the following result ([Ste93, X.2.3]): if $\mathcal{R}$ is the family of rectangles in $\mathbb{R}^2$, centered at the origin and with sides parallel to the axes, there exists $f \in L^1(\mathbb{R}^2)$ such that

$$\limsup_{\operatorname{diam} R\to 0} \frac{1}{m(R)} \int_{x+R} f = \infty \qquad \text{a.e. } (x)$$

where $R$ moves in $\mathcal{R}$.

Another consequence of Lebesgue differentiation is the following generalization of the Fundamental Theorem of Calculus

**2.11.14. Corollary.** *(Fundamental Theorem of Calculus)*

*Let $f \in L^1(\mathbb{R})$, and set $F(x) = \int_{-\infty}^{x} f\, dm$. Then $F$ is differentiable a.e. and $F' = f$ a.e. More concretely, the equality holds at every Lebesgue point of $f$.*

*Proof.* Let $E_k = [x, x+\delta_k]$, where $\delta_k > 0$ and $\delta_k \to 0$. These sets trivially shrink nicely to $x$. Indeed, taking $r_k = 2\delta_k$ and $\alpha = 1/5$ we have $E_k \subset B_{r_k}(x) = (x - 2\delta, x + 2\delta) \supset E_k$ and $m(E_k) = \delta_k > \alpha\, 4\delta = \alpha\, m(B_{r_k}(x))$. Then, if $x$ is a Lebesgue point for $f$,

$$\frac{F(x+\delta_k) - F(x)}{\delta_k} = \frac{1}{m(E_k)} \int_{E_k} f\, dm \xrightarrow[k\to\infty]{} f(x).$$

As we are free to choose the sequence $\{\delta_k\}$, this shows that the right-derivative of $F$ at $x$ exists and is equal to $f(x)$. A similar argument shows likewise for the left derivative. So $F$ is differentiable at $x$ and $F'(x) = f(x)$. $\square$

When $f'$ is continuous, one obtains the most common use of the Fundamental Theorem of Calculus,

$$f(b) - f(a) = \int_a^b f'\, dm, \tag{2.66}$$

often called **Barrow's Rule**. One runs into problems if $f'$ is not integrable. A typical example is given by

$$f(x) = x^2 \sin\frac{1}{x^2}, \qquad x \in \mathbb{R} \setminus \{0\},$$

with $f(0) = 0$. This $f$ is differentiable everywhere; the only non-obvious point is $x = 0$, where we get that $f'(0) = 0$ via squeezing with parabolas. But

$$f'(x) = 2x \sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2}$$

is not continuous at 0, and $\int_0^x f'\, dm$ does not exist as a Lebesgue integral, though it does exist as an improper Riemann integral; concretely, it is the second term that causes problems. For the first term, both the positive and negative parts are continuous and bounded, thus integrable. But for the second term,

$$\int_0^1 \frac{2}{x}\cos^+\frac{1}{x^2}\, dm \geq \sum_{k=1}^{\infty} \int_{\frac{1}{\sqrt{(2k+\frac14)\pi}}}^{\frac{1}{\sqrt{(2k-\frac14)\pi}}} \frac{2}{x}\cos\frac{1}{x^2}\, dm \geq \sum_{k=1}^{\infty} \int_{\frac{1}{\sqrt{(2k+\frac14)\pi}}}^{\frac{1}{\sqrt{(2k-\frac14)\pi}}} \frac{2}{\frac{1}{\sqrt{(2k+\frac14)\pi}}}\, \frac{1}{\sqrt{2}}$$

$$
\begin{aligned}
&= \sqrt{2}\sum_{k=1}^{\infty}\sqrt{2k+\frac{1}{4}}\Big(\frac{1}{\sqrt{2k-\frac{1}{4}}}-\frac{1}{\sqrt{(2k+\frac{1}{4})}}\Big)\\
&= \frac{\sqrt{2}}{2}\sum_{k=1}^{\infty}\frac{1}{\sqrt{2k-\frac{1}{4}}}\frac{1}{\sqrt{2k+\frac{1}{4}}+\sqrt{2k-\frac{1}{4}}}\\
&\geq \frac{\sqrt{2}}{4}\sum_{k=1}^{\infty}\frac{1}{\sqrt{2k+\frac{1}{4}}}=\infty.
\end{aligned}
$$

A similar computation shows that $(f')^-$ is not integrable either.

Another problem with (2.66) occurs when $f$ is differentiable a.e., even if $f' \in L^1(\mathbb{R})$. For instance consider the Cantor Function $\alpha$ from page 87. Then $\alpha' = 0$ a.e. while $\alpha(1) - \alpha(0) = 1$.

If we consider $d\mu = f'\,dm$ in (2.66), then $\mu \ll m$. By Propositions 2.10.6 and 2.10.7, for each $\varepsilon > 0$ there exists $\delta > 0$ such that $|\mu|(E) < \varepsilon$ whenever $m(E) < \delta$. Using the definition of the total variation we can rephrase this as saying that for any $\varepsilon > 0$ and any $E$ with $m(E) < \delta$ and any partition $\{E_k\}$ of $E$,

$$\sum_k |\mu(E_k)| < \varepsilon.$$

As $\mu((a,b)) = f(b) - f(a)$, this suggests the following definition.

**2.11.15. Definition.**

*A function $f : [a,b] \to \mathbb{C}$ is **absolutely continuous** on $[a,b]$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\sum_{j=1}^{n} |f(b_j) - f(a_j)| < \varepsilon$$

*whenever $a \leq a_1 < b_1 < a_2 < b_2 < \cdots < b_n \leq b$ and $\sum_j |b_j - a_j| < \delta$.*

*The **Total Variation** of $f$ on $[a,b]$ is*

$$F(x) = \sup\Big\{\sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| : \; a = t_0 < t_1 < \cdots < t_N = x\Big\}.$$

*When $F(b) < \infty$ we say that $f$ is of **bounded variation**.*

Functions of bounded variation are fairly nice. For instance, their side limits always exist (Exercise 2.11.7), their real and imaginary parts can always be written as the difference of two monotone functions (Exercise 2.11.8), and they are always Riemann integrable (Exercise 2.11.9).

Using the case where $n = 1$ we see that absolute continuity implies continuity. It turns out that not only is absolute continuity a necessary condition for

Barrow's Rule (2.66), it is also sufficient. Towards that goal, we first show two results.

***2.11.16. Proposition.***

*Let $f : [a,b] \to \mathbb{R}$ be absolutely continuous and $F$ its total variation. Then $F, F+f$ and $F-f$ are non-decreasing and absolutely continuous.*

*Proof.* If $a < x < y < b$ and $a \leq t_0 < \cdots < t_N \leq x$, then we can see these points as a partition of $[a,x]$ and these points together with $y$ as a partition of $[a,y]$; by definition of the total variation we then have

$$F(y) \geq |f(y) - f(x)| + \sum_{j=1}^{N} |f(t_j) - f(t_{j-1})|.$$

As we can do this for any such choice of $t_0, \ldots, t_N$ we get that $F(y) \geq |f(y) - f(x)| + F(x)$. Thus $F(y) \geq F(x)$ and so $F$ is non-decreasing. We also get $F(y) \geq f(y) - f(x) + F(x)$, and so $F(y) - f(y) \geq F(x) - f(x)$; and from $F(y) \geq f(x) - f(y) + F(x)$ we also obtain $F(y) + f(y) \geq F(x) + f(x)$. Finally, a straightforward application of the triangle inequality implies that linear combinations of absolutely continuous functions are again absolutely continuous (Exercise 2.10.15). □

***2.11.17. Proposition.***

*Let $f : [a,b] \to \mathbb{R}$ continuous and non-decreasing. The following statements are equivalent:*

*(i) $f$ is absolutely continuous on $[a,b]$;*

*(ii) $f$ maps nullsets to nullsets;*

*(iii) $f$ is differentiable a.e. on $[a,b]$, $f' \in L^1(\mathbb{R})$, and $f(x) - f(a) = \int_a^x f' \, dm$ for all $x \in [a,b]$.*

*Proof.* (i) $\implies$ (ii) Let $E \subset [a,b]$ with $m(E) = 0$. Since $f$ maps points to nullsets, we may remove $a, b$ from $E$ if they were present. Fix $\varepsilon > 0$ and let $\delta > 0$ be the one provided by the absolute continuity of $f$. By Proposition 2.3.25 there exists $V$ open with $E \subset V \subset [a,b]$ and $m(V) < \delta$. Writing $V = \bigcup_j (a_j, b_j)$ (Proposition 1.8.1), we have $\sum_j b_j - a_j < \delta$ and so

$$\sum_j f(g_j) - f(a_j) < \varepsilon.$$

As $E \subset V$ and $f$ is continuous and non-decreasing (so it maps intervals to intervals)

$$f(E) \subset f(V) \subset \bigcup_j [f(a_j), f(b_j)].$$

Thus $m^*(f(E)) < \varepsilon$ for all $\varepsilon > 0$, so $m^*(f(E)) = 0$, showing that $f(E) \in \mathcal{M}(\mathbb{R})$ and $m(f(E)) = 0$.

(ii) $\implies$ (iii) Let $g(x) = x + f(x)$. If $E$ is a nullset, $m^*(g(E)) \leq m^*(E) + m^*(f(E)) = 0$. So $g$ maps nullsets to nullsets. Let $E \subset [a, b]$ be measurable. By Proposition 2.3.28 we can write $E = E_1 \cup E_0$ with $E_1$ a countable union of closed, and $E_0$ a nullset. Since $[a, b]$ is compact, the set $E_1$ is a countable union of compact. By the continuity of $g$, and since continuous functions map compact sets to compact set, $g(E_1)$ is a countable union of compact. And as $g$ maps nullsets to nullsets, $g(E_0)$ is a nullset. Then $g(E) = g(E_1) \cup g(E_0) \in \mathcal{M}(\mathbb{R})$. Define a measure $\mu$ on the measurable subsets of $[a, b]$ by $\mu(E) = m(g(E))$. The $\sigma$-additivity can be checked from the fact that $g$ is strictly increasing. Since $\mu \ll m$ (because $g$ maps nullsets to nullsets and $\mu$ is bounded, by Theorem 2.10.10 there exists $h \in L^1(\mathbb{R})$ such that $d\mu = h\, dm$.

If we take $E = [a, x]$, then $g(E) = [g(a), g(x)]$ and

$$g(x) - g(a) = m(g(E)) = \mu(E) = \int_E h\, dm = \int_a^x h\, dm.$$

Then

$$f(x) - f(a) = \int_a^x (h - 1)\, dm.$$

As $h - 1 \in L^1(\mathbb{R})$, by Corollary 2.11.14 we get that $f$ is differentiable a.e. and $f' = h - 1$ a.e.

(iii) $\implies$ (i) For a partition of $[a, b]$ we have

$$f(b_j) - f(a_j) = \int_a^{b_j} f' - \int_a^{a_j} f' = \int_{a_j}^{b_j} f'.$$

Then

$$\sum_{j=1}^n |f(b_j) - f(a_j)| = \sum_{j=1}^n \left| \int_{a_j}^{b_j} f' \right| \leq \sum_{j=1}^n \int_{a_j}^{b_j} |f'| = \int_{E_n} |f'|$$

where $E_n = \bigcup_{j=1}^n [a_j, b_j]$. The result now follows from Exercise 2.8.10. $\square$

We can now put everything together:

**2.11.18. Theorem.**

*Let $f : [a,b] \to \mathbb{C}$ be absolutely continuous. Then $f$ is differentiable a.e., $f' \in L^1[a,b]$, and*

$$f(x) - f(a) = \int_a^x f'\, dm, \qquad a \le x \le b.$$

*Proof.* Assume first that $f$ is real. By Proposition 2.11.16 we know that $g_1 = \frac{1}{2}(F+f)$ and $g_2 = \frac{1}{2}(F-f)$ are non-decreasing and absolutely continuous. By Proposition 2.11.17 both functions are differentiable a.e. and satisfy Barrow's Rule (2.66). Then

$$\begin{aligned} f(x) - f(a) = g_1(x) - g_1(a) - [g_2(x) - g_2(a)] &= \int_a^x g_1'\, dm - \int_a^x g_2'\, dm \\ &= \int_a^x (g_1 - g_2)'\, dm = \int_a^x f'\, dm. \end{aligned}$$

We also have

$$\int_a^b |f'|\, dm \le \int_a^b (g_1' + g_2')\, dm = \int_a^b F'\, dm = F(b) - F(a) < \infty$$

and so $f' \in L^1[a,b]$. □

**2.11.1. Exercises.**

(2.11.1) Consider $\mathbb{R}$ with Lebesgue measure, and $E \subset \mathbb{R}$ measurable. If it exists, the number

$$d_E(x) = \lim_{\varepsilon \to 0} \frac{m(E \cap (x-\varepsilon, x+\varepsilon))}{2\varepsilon}$$

is the *density* of $E$ at $x$. Show that $d_E(x) = 1_E(x)$ a.e. Can you formulate and prove an analog result in $\mathbb{R}^n$?

(2.11.2) Let $f \in L^1[0,\infty)$ be such that

$$\int_0^x f\, dm = 0, \qquad x > 0.$$

Show that $f = 0$. This was already done in Exercise 2.5.6 but now a much much shorter proof is available.

(2.11.3) For a closed square and a closed disk in $\mathbb{R}^2$, calculate the density at every point.

(2.11.4) For $E \subset \mathbb{R}^2$, the boundary $\partial E$ is the closure of $E$ minus the interior of $E$.

(i) Show that $E$ is Lebesgue measurable if $m(\partial E) = 0$.

(ii) Suppose that $E$ is an arbitrary union of a collection of closed disks with radii at least $c$ for some $c > 0$. Show that $E$ is measurable.

(iii) Show that the radii above don't need to be restricted.

(iv) Show that some unions of closed disks of radius 1 are not Borel sets.

(v) Can disks be replaced by triangles, rectangles, arbitrary polygons, etc.? What is the relevant geometric property?

(2.11.5) Let $\{f_n\}$ be a sequence of non-decreasing functions $f_n : \mathbb{R} \to [0, \infty)$, such that $f(x) = \sum_n f_n(x) < \infty$ for all $x$. Show that $f'(x) = \sum_n f_n'(x)$ a.e.

(2.11.6) Suppose that $E \subset [a, b]$, $m(E) = 0$. Construct an absolutely continuous monotonic function $f$ on $[a, b]$ so that $f'(x) = \infty$ for all $x \in E$.

(2.11.7) Let $f : [a, b] \to \mathbb{C}$ be of bounded variation. Show that $f$ admits side-limits at all points.

(2.11.8) Let $f : [a, b] \to \mathbb{R}$. Show that $f$ is of bounded variation if and only if there exist $g, h : [a, b] \to \mathbb{R}$, both monotone, and such that $f = g - h$.

(2.11.9) Let $f : [a, b] \to \mathbb{C}$ be a function of bounded variation. Show that $f$ is Riemann integrable.

CHAPTER

# 3 A Bit of Complex Analysis

Complex Analysis underlies several key results in Functional Analysis, particularly in Chapter 9. The idea in this chapter is to provide some technical detail and the main results of the theory. Because of this we will not be looking for maximum generality. In any case, complex Analysis is a beautiful theory that deserves to be known, and we hope the reader will enjoy it in this chapter.

## 3.1. Analytic and Holomorphic Functions

Besides the usual $\mathbb{C}$ for the complex plane, we denote by $\mathbb{D}$ the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : \ |z| < 1\}$ and by $\mathbb{T}$ the unit circle $\mathbb{T} = \{z \in \mathbb{C} : \ |z| = 1\}$.

Let us begin with a basic result about series.

***3.1.1. Lemma. (The Ratio Test)***

*Let $\sum_{k=0}^{\infty} b_k$ be a series of nonzero complex numbers. Let*

$$\alpha = \limsup_{k\to\infty} \frac{|b_{k+1}|}{|b_k|}, \qquad \beta = \liminf_{k\to\infty} \frac{|b_{k+1}|}{|b_k|}.$$

*Then*

*(i) if $\alpha < 1$, the series converges absolutely;*

*(ii) if $\beta > 1$, the series is divergent.*

*Proof.* Suppose $\alpha < 1$. Then there exists $m \in \mathbb{N}$ and $\delta > 0$ such that $\frac{|b_{k+1}|}{|b_k|} < 1-\delta$ for all $k \geq m$. It follows inductively that $|b_{m+k}| \leq (1-\delta)^k |b_m|$. Then

$$\sum_{k>m} |b_k| \leq \sum_{k=1}^{\infty} (1-\delta)^k |b_m| = \frac{(1-\delta)|b_m|}{\delta} < \infty$$

and the series converges absolutely.

When $\beta > 1$, we now have that there exists $m \in \mathbb{N}$ and $\delta > 0$ such that $|b_{k+1}| > (1+\delta)|b_k|$ for all $k \geq m$. Then $|b_{m+k}| \geq (1+\delta)^k |b_m|$, showing that $b_k$ does not converge to zero, and so the series diverges. □

It is well-known that when $\alpha = 1$ or $\beta = 1$, no information can be obtained about the series. Concretely, the divergent series $\sum_k \frac{1}{k}$ has $\alpha = \beta = 1$, while the convergent series $\sum_k \frac{1}{k^2}$ also has $\alpha = \beta = 1$. A variation of this idea can produce both convergent and divergent series of positive terms with $\alpha > 1$ and $\beta < 1$ (Exercise 3.1.1).

Now the basic definitions. We will repeatedly denote by $z_0 + r\,\mathbb{D}$ the open disk of radius $r$ around $z_0$.

**3.1.2. Definition.**

*Let $r > 0$, $z_0 \in \mathbb{C}$. We say that a function $f : z_0 + r\,\mathbb{D} \to \mathbb{C}$ is **analytic** at $z_0$ if there exist $\{a_n\}_{n\in\mathbb{N}\cup\{0\}} \subset \mathbb{C}$ such that*

$$f(z) = \sum_{k=0}^{\infty} a_k\,(z - z_0)^k, \qquad z \in z_0 + r\,\mathbb{D}.$$

*Let $V \subset \mathbb{C}$ be open and $f : V \to \mathbb{C}$. We say that $f$ is **holomorphic** on $V$ if the derivative*

$$f'(z) = \lim_{w\to z} \frac{f(w) - f(z)}{w - z}$$

*exists for all $z \in V$. When $f$ is holomorphic on all of $\mathbb{C}$, we say that $f$ is **entire**.*

Easy examples of holomorphic functions are the polynomials and rational functions as long as the domain excludes the zeros of the denominator. Easy examples of analytic functions are the polynomials and the exponential, sine, and cosine, namely

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}, \quad \sin z = \sum_{k=1}^{\infty} \frac{(-1)^k z^{2k-1}}{(2k-1)!}, \quad \cos z = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!}.$$

It is also useful to note that

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \qquad \cos z = \frac{e^{iz} + e^{-iz}}{2}.$$

Also, $\frac{d}{dz} e^z = e^z \neq 0$ for all $z$. But the exponential is not bijective over the complex plane, as it is periodic: concretely, $e^{2k\pi i} = 1$ for all $k \in \mathbb{Z}$.

By basic properties of series, the sum and the product of power series is a power series. The same is true about holomorphic functions, this time by basic properties of derivatives; that is, sums, products, and compositions of holomorphic functions are holomorphic, provided that the domains match. Other than than, holomorphic functions look a bit mysterious, but that will change soon.

Let us first gather some information about analytic functions. The **radius of convergence** of the analytic function $f(z) = \sum_{k=0}^{\infty} a_k\,(z - z_0)^k$ is the number (possibly infinity)

$$R[f] = \sup\{s \geq 0 : \sum_k |a_k|\, s^k < \infty\}.$$

***3.1.3. Proposition.***

*Let $r > 0$ and $f : z_0 + r\,\mathbb{D} \to \mathbb{C}$ be analytic at $z_0$, with $f(z) = \sum_{k=0}^{\infty} a_k\,(z - z_0)^k$. Then*

*(i) the series converges absolutely in $z_0 + r\,\mathbb{D}$ and uniformly in $z_0 + s\,\mathbb{D}$ for any $s \in (0, r)$;*

*(ii) $f$ is infinitely differentiable; in particular, it is holomorphic;*

*(iii) the coefficients $\{a_k\}$ are unique; more specifically,*

$$a_k = \frac{f^{(k)}(z_0)}{k!}, \qquad k \in \mathbb{N} \cup \{0\}; \tag{3.1}$$

*(iv) $R[f] = \liminf_n |a_n|^{-1/n}$;*

*(v) the series converges absolutely when $|z - z_0| < R[f]$, and it diverges when $|z - z_0| > R[f]$;*

*(vi) $r \leq R[f]$;*

*(vii) for all $k \in \mathbb{N}$, $R[f^{(k)}] = R[f]$.*

*Proof.*

(i) Fix $z \in z_0 + r\,\mathbb{D}$. Because $\mathbb{D}$ is open, there exists $z_1 \in z_0 + r\,\mathbb{D}$ with $|z_1 - z_0| > |z - z_0|$. As the series $\sum_k a_k\,(z_1 - z_0)^k$ converges, we have that $\lim_k a_k(z_1 - z_0)^k = 0$. So there exists $k_0$ such that for all $k > k_0$, $|a_k|\,|z_1 - z_0|^k < 1$. Fix $s$ with $|z_1 - z_0| > s \geq |z - z_0|$. For any $k_1 > k_0$,

$$\sum_{k\geq k_1} |a_k|\,|z-z_0|^k \leq \sum_{k\geq k_1} \left(\frac{s}{|z_1-z_0|}\right)^k.$$

As the bound goes to zero and it does not depend on $z$, we get that the series converges absolutely and uniformly.

(ii) Fix $z \in z_0 + r\,\mathbb{D}$. Choose $s$ with $|z-z_0| < s < r$. Then the series converges uniformly on $z_0 + s\,\mathbb{D}$; and the series of the derivatives has terms $k\,a_k z^{k-1}$, so as above

$$\sum_{k\geq k_1} k|a_k|\,|z-z_0|^{k-1} \leq \frac{1}{s}\sum_{k\geq k_1} k\left(\frac{s}{|z_1-z_0|}\right)^k.$$

This last series converges by Lemma 3.1.1, since

$$\frac{(k+1)\left(\frac{s}{|z_1-z_0|}\right)^{k+1}}{k\left(\frac{s}{|z_1-z_0|}\right)^{k}} = \frac{k+1}{k}\,\frac{s}{|z_1-z_0|} \xrightarrow[k\to\infty]{} \frac{s}{|z_1-z_0|} < 1.$$

So the series of the derivatives converges uniformly on $z_0 + s\overline{\mathbb{D}}$.

Fix $z \in z_0 + s\,\mathbb{D}$. Choose $s_0$ with $|z-z_0| < s_0 < s$. Then, for $h \in \mathbb{C}$ with $|h| < (s-s_0)/2$, and writing $w = z - z_0$,

$$\begin{aligned}
\Big|\frac{f(z+h)-f(z)}{h} - \sum_{k=1}^{\infty} k\,a_k\,w^{k-1}\Big| &= \Big|\sum_{k=2}^{\infty} a_k\,\frac{(w+h)^k - w^k - hkw^{k-1}}{h}\Big| \\
&= \Big|\sum_{k=2}^{\infty}\frac{1}{h}\,a_k\sum_{j=1}^{k}\binom{k}{j}h^j w^{k-j} - hkw^{k-1}\Big| \\
&= \Big|\sum_{k=2}^{\infty}\frac{1}{h}\,a_k\sum_{j=2}^{k}\binom{k}{j}h^j w^{k-j}\Big| \\
&\leq \sum_{k=2}^{\infty}|a_k|\sum_{j=2}^{k}\binom{k}{j}|h|^{j-1}s_0^{k-j} \\
&= |h|\sum_{k=2}^{\infty}|a_k|\sum_{j=0}^{k-2}\binom{k}{j+2}|h|^j s_0^{k-2-j} \\
&\leq |h|\sum_{k=2}^{\infty}k(k-1)|a_k|\sum_{j=0}^{k-2}\binom{k-2}{j}|h|^j s_0^{k-2-j} \\
&= |h|\sum_{k=2}^{\infty}k(k-1)|a_k|(s_0+|h|)^{k-2}.
\end{aligned}$$

This last series converges because, as above, the series $\sum_k a_k s^k$ converges, which gives us that eventually $|a_k| < s^{-k}$ and $s_0 + |h| < s$ and we can apply the Lemma 3.1.1. Hence

$$f'(z) = \sum_{k=1}^{\infty} k\, a_k\, (z - z_0)^{k-1} = \sum_{k=0}^{\infty} (k+1)\, a_{k+1}\, (z - z_0)^k.$$

As this is a power series that converges in all of $z_0 + r\,\mathbb{D}$, everything done in the proof so far applies, so $f'$ is differentiable, and we can iterate the process.

(iii) We proceed by induction. For $k = 0$, $a_0 = f(z_0) = \frac{f^{(0)}(z_0)}{0!}$. Assume that, for any power series $\sum_k a_k\, (z - z_0)^k$, the equality (3.1) holds. In particular it applies to the series for $f'(z) = \sum_{k=0}^{\infty} (k+1)\, a_{k+1}\, (z - z_0)^k$. Applying (3.1) to the $k^{\text{th}}$ coefficient of $f'$, we obtain

$$(k+1)\, a_{k+1} = \frac{(f')^{(k)}(z_0)}{k!} = \frac{f^{(k+1)}(z_0)}{k!},$$

showing that

$$a_{k+1} = \frac{f^{(k+1)}(z_0)}{(k+1)!}$$

(iv) Let $t = \liminf_n |a_n|^{-1/n}$. If $0 < s < t$, choose $\delta > 0$ with $t - \delta > s$. For $n$ big enough, say $n > n_0$, $|a_n|^{-1/n} > t - \delta$, which we may write as $|a_n| < (t - \delta)^{-n}$. Then

$$\sum_{n > n_0} |a_n|\, s^n \le \sum_{n > n_0} \left(\frac{s}{t - \delta}\right)^n < \infty.$$

Thus $t \le R[f]$. Now if $s > t$, choose $\delta > 0$ with $s > t + \delta$. Then there is a subsequence $\{a_{n_k}\}$ with $|a_{n_k}|^{-1/n_k} < t + \delta$; we may write this as $|a_{n_k}| > (t + \delta)^n$. Then

$$a_{n_k} s^{n_k} > \left(\frac{s}{t + \delta}\right)^n > 1$$

for all $k$ and the series cannot converge. So $R[f] \le t$ and thus $R[f] = t$.

(v) This is now a restatement of what we just proved above.

(vi) By what we have proven, $\sum_k |a_k| s^k$ converges for all $s < r$. Then $r \le R[f]$

(vii) We have, since $k^{-1/k} = e^{\frac{1}{k}\,\log k} \to 1$,

$$R[f'] = \liminf_{k \to \infty} k^{-1/k} |a_k|^{-1/k} = \liminf_{k \to \infty} |a_k|^{-1/k} = R[f].$$

For higher derivatives, we simply iterate the equality. $\square$

It is tempting to think that in Proposition 3.1.3 one should get that $R[f] = r$. But that cannot possibly be the case in general, because the radius $r$ might have been imposed artificially, and there is not way for the coefficients $\{a_k\}$ to know about $r$.

### 3.1.1. Exercises.

(3.1.1) Construct an example of a convergent series of positive terms $\sum_k b_k$ such that $\alpha > 1$ and $\beta < 1$.

(3.1.2) Construct an example of a divergent series of positive terms $\sum_k b_k$ such that $\alpha > 1$ and $\beta < 1$.

(3.1.3) Show that the function $f(z) = \bar{z}$ is not holomorphic anywhere in the complex plane.

(3.1.4) Prove the **Dirichlet Criterion**: If $\{a_n\}, \{b_n\} \subset \mathbb{C}$ and

(i) $\lim\limits_{n\to\infty} a_n = 0$;

(ii) $\sum\limits_{n=1}^{\infty} |a_{n+1} - a_n| < \infty$;

(iii) there exists $M > 0$ such that $\left|\sum\limits_{n=1}^{m} b_n\right| < M$ for all $m$.

Show that $\sum\limits_{n} a_n b_n$ converges.

(3.1.5) Show that in Exercise 3.1.4 the hypothesis "$\sum\limits_{n=1}^{\infty} |a_{n+1} - a_n| < \infty$" can be replaced, when it makes sense, by "$\{a_n\}$ is monotone".

(3.1.6) Show that the **Gamma Function**

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

defines a holomorphic function on the semiplane $\operatorname{Re} z > 0$. Show also that $\Gamma(z+1) = z\Gamma(z)$ for every $z$ in its domain, and that $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$.

## 3.2. Inverses of Holomorphic Functions and the Logarithm

Let us now turn our focus to holomorphic functions for a while. We have already shown that every analytic function is holomorphic. As already mentioned, the polynomials, the exponential, and the basic trigonometric functions are holomorphic—at least locally for those trigonometric functions defined as quotients, like the tangent.

**3.2.1. Theorem. (Inverse Function Theorem)**

*Let $V, W \subset \mathbb{C}$ be open, $f : V \to \mathbb{C}$, $g : W \to \mathbb{C}$, with $f(V) \subset W$ and $g(f(z)) = z$ for all $z \in V$. If $g$ is holomorphic at $f(z_0)$ and $g'(f(z_0)) \neq 0$, then $f$ is holomorphic at $z_0$ and*

$$f'(z_0) = \frac{1}{g'(f(z_0))}. \tag{3.2}$$

*Proof.* For any nonzero $h \in \mathbb{C}$ such that $z_0 + h \in V$. We have $g(f(z_0+h)) = z_0 + h \neq z_0 = g(f(z_0))$, so $f(z_0+h) \neq f(z_0)$. Since $f(z_0+h) - f(z_0) \to 0$ with $h$ by the continuity of $f$,

$$\lim_{h\to 0} \frac{g(f(z_0+h)) - g(f(z_0))}{f(z_0+h) - f(z_0)} = g'(f(z_0)) \neq 0.$$

Then

$$\lim_{h\to 0} \frac{f(z_0+h) - f(z_0)}{h} = \lim_{h\to 0} \frac{1}{\frac{g(f(z_0+h)) - g(f(z_0))}{f(z_0+h) - f(z_0)}} = \frac{1}{g'(f(z_0))}.$$

So $f$ is holomorphic at $z_0$, and (3.2) is established. □

The logarithm function is already a bit quirky on the real line; and things are worse in the complex plane. What the logarithm should look like, is not hard to figure. Indeed, if $z = re^{i\theta}$ and we want to write $z = e^w$, we need $w = \log r + i\theta$. This makes us write

$$\log re^{i\theta} = \log r + i\theta. \tag{3.3}$$

Since $\exp(\log z) = z$, we expect from Theorem 3.2.1 that $\log z$ be holomorphic. The issue is that we cannot define log globally on all of $\mathbb{C}$. It turns out that there is a problem with (3.3); consider the complex number $w = 2e^{17i/3}$. According to (3.3), we want to write $\log w = \log 2 + i\,\frac{17}{3}$. But we also have $w = 2e^{5i/3}$, which would make us write $\log w = \log 2 + i\,\frac{5}{3}$. So (3.3) is not really well-defined. Some authors prefer to say that the logarithm function is *multivalued*. On a first look one could be tempted to say, "well, that's no big deal; we just take the argument $\theta$ to be in $[0, 2\pi)$ and the problem above does not appear". That is partially true, but there is still a problem. Consider the circle $\gamma(t) = 2e^{it}$, $t \in [0, 2\pi]$. If we apply the logarithm as defined above, we get $\log(\gamma(t)) = \log 2 + it$. As we finish going around the circle, we approach the starting point $\gamma(0) = \gamma(2\pi) = 2$; but now the values of the logarithm are approaching $\log 2 + \pi i$, so $\log \circ \gamma$ is not continuous. This is something we want to avoid, as we are dealing with holomorphic functions which are by force continuous.

Choosing one interval for the argument of the logarithm is usually referred to as "choosing a branch". The canonical choice, called the **principal branch** of the logarithm is the choice $\theta \in (-\pi, \pi)$. This removes the negative axis from the domain, and so we obtain—by Theorem 3.2.1—a holomorphic function $z \longmapsto \log z$ on $\mathbb{C}\backslash(-\infty, 0]$. At times, we will use other branches than the principal one.

A consequence of this feature of the logarithm function, is that while we have $\exp(\log z) = z$ for all $z$ in the domain of whatever branch we are using, it is not true in general that $\log(\exp(z)) = z$; this equality will only hold when the argument of $z$ lies within the interval of choice of the arguments. For example, with the principal branch we have

$$\log\left(\exp\left(3 + \frac{7i\pi}{3}\right)\right) = \log(e^3\, e^{7i\pi/3}) = \log(e^3\, e^{i\pi/3}) = 3 + \tfrac{i\pi}{3}.$$

After all the above, whenever we fix a region that does not contain a fixed line that starts at the origin we can define a branch $\log z$ of the logarithm. This function is continuous, since the choice of the branch guarantees that the argument cannot "turn around"; so points that are close in the domain of $\log z$ (in the sense of being joined by a short path) will have close arguments. Then Theorem 3.2.1, applied to $\exp(\log z) = z$, gives us that $\log z$ is holomorphic and that

$$(\log z)' = \frac{1}{z}.$$

## 3.3. Line Integrals

A crucial aspect of Complex Analysis has to do with line integrals and the behaviour of holomorphic functions with respect to them.

**3.3.1. Definition.**

*We say that a function $\gamma : [a, b] \to \mathbb{C}$ is **piecewise continuously differentiable** if $\gamma$ if there exists a partition $a = a_0 < a_1 < \cdots < a_n = b$ such that $\gamma$ is continuously differentiable on $[a_k, a_{k+1}]$ for all $k = 0, \ldots, n-1$. We will often refer to such $\gamma$ simply as a **curve**.*

*Given a piecewise continuously differentiable curve $\gamma : [a, b] \to \mathbb{C}$ and a complex function $f$, we define the **integral of $f$ along** $\gamma$ as*

$$\int_\gamma f(z)\, dz = \int_a^b f(\gamma(t))\, \gamma'(t)\, dt. \tag{3.4}$$

The equality (3.4) certainly looks familiar[1]. It kind of hides the fact that we are dealing with complex functions. But that introduces no complication, other than in terms of calculus there are two equalities in (3.4), one for the real part and one for the imaginary part. Or one notices that derivatives and integrals behave the same when we work with complex instead of real scalars.

It is important to note that the definition of piecewise continuous differentiability we are using implies that $\gamma'$ is bounded on $[a,b]$. This implies that $\gamma$ is of bounded variation (Exercise 3.3.1).

Since we allow piecewise continuity, it will be convenient to use a different domain for each continuous piece of $\gamma$. If we denote the joining of two continuous pieces $\gamma_1$ and $\gamma_2$ by $\gamma_1+\gamma_2$, we can say that a curve in general is of the form $\gamma=\sum_{k=1}^m \gamma_k$, with each $\gamma_k:[a_k,b_k]\to\mathbb{C}$ continuously differentiable. It is common to relax this condition a bit and only require that $\gamma$ is **rectifiable**, which means that $\gamma$ is piecewise continuous and that its length is well-defined. All the curves we will deal with are piecewise differentiable; but one should keep in mind that all results apply to rectifiable curves.

For a general curve $\gamma=\sum_k \gamma_k$ as above, we define

$$\int_\gamma f(z)\,dz=\sum_{k=1}^m \int_{\gamma_k} f(z)\,dz.$$

***Important:*** *every time we refer to a "curve" or "path" $\gamma$, we mean a piecewise continuously differentiable curve as above. We also use $\gamma$ to denote the set*

$$\gamma=\bigcup_k \gamma_k([a_k,b_k]).$$

*We also write $\gamma=\gamma_1+\cdots+\gamma_m$. And $\gamma_1-\gamma_2$ means $\gamma_1+(-\gamma_2)$, where $-\gamma_2$ is the reverse curve, i.e., $-\gamma_2(t)=\gamma_2(a_k+b_k-t)$, $t\in[a_k,b_k]$. It is straightforward to check that*

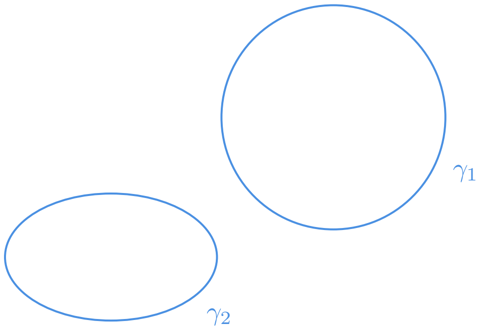
$$\int_{\gamma_1-\gamma_2} f(z)\,dz=\int_{\gamma_1} f(z)\,dz-\int_{\gamma_2} f(z)\,dz.$$

A continuous curve $\gamma:[a,b]\to\mathbb{C}$ is **closed** if $\gamma(a)=\gamma(b)$. When $\gamma$ is not continuous but it has several continuous components, we will say that $\sum_k \gamma_k$ is closed if each $\gamma_k$ is closed. So for instance two disjoint circles form a closed curve. For integrals over closed curves, it is common to use the notation

$$\oint_\gamma f(z)\,dz,$$

to emphasize the fact that the curve is closed.

[1]unless you got here without knowing calculus; that would be unusual enough that if such is the case I beg you to contact me and tell me how it happened!




*A closed curve* $\gamma = \gamma_1 + \gamma_2$.

Being defined in terms of Riemann integration, this line integral behaves in the usual, nice way. For instance,

**3.3.2. Proposition.**

*(i)* $\displaystyle\int_\gamma (f(z) + \lambda\, g(z))\, dz = \int_\gamma f(z)\, dz + \lambda \int_\gamma g(z)\, dz$ *for all* $\lambda \in \mathbb{C}$, *and* $f, g$ *integrable.*

*(ii)*
$$\int_{-\gamma} f(z)\, dz = -\int_\gamma f(z)\, dz.$$

*(iii)* $\displaystyle\int_{\gamma_1+\gamma_2} f(z)\, dz = \int_{\gamma_1} f(z)\, dz + \int_{\gamma_2} f(z)\, dz.$

Our initial goal, that will open the doors to understanding holomorphic functions completely, is to show that that every holomorphic function has an antiderivative. This is trivial to do in calculus, over the real line; here, it will require a bit of work.

Here is an estimate that will prove useful later.

**3.3.3. Lemma. (Jordan's Lemma)**

*Let* $\gamma(t) = Re^{it}$, $t \in [0, \pi]$, $\beta > 0$, *and* $g : \mathbb{C} \to \mathbb{C}$ *continuous. Then*
$$\left| \int_\gamma e^{i\beta z} g(z)\, dz \right| \leq \frac{\pi \max_\gamma |g|}{\beta}.$$

*Proof.* By definition,
$$\int_\gamma e^{i\beta z} g(z)\, dz = \int_0^\pi g(Re^{it})\, e^{i\beta\, R(\cos t + i \sin t)}\, Ri\, e^{it}\, dt$$

$$= \int_0^{\pi} g(Re^{it})\, e^{i\beta\, R\,\cos t}\, e^{-\beta R\,\sin t}\, Ri\, e^{it}\, dt.$$

Then, with $M = \max_\gamma |g|$,

$$\Big|\int_\gamma e^{i\beta z} g(z)\, dz\Big| \le R \int_0^{\pi} |g(R\, e^{it})|\, e^{-\beta R\,\sin t}\, dt$$

$$\le MR \int_0^{\pi} e^{-\beta R\,\sin t}\, dt = 2MR \int_0^{\pi/2} e^{-\beta R\,\sin t}\, dt.$$

The last equality comes from $\int_{\pi/2}^{\pi} f(\sin t)\, dt = \int_0^{\pi/2} f(\sin t)\, dt$ via substitution. The function $t \longmapsto \frac{\pi}{2}\sin t$ maps $[0, \frac{\pi}{2}]$ onto itself. As it is concave, we get $\frac{\pi}{2}\sin t \ge t$ for all $t \in [0, \frac{\pi}{2}]$. As $\beta$ and $R$ are positive,

$$\left|\int_\gamma e^{i\beta z} g(z)\, dz\right| \le 2MR \int_0^{\pi/2} e^{-2t\beta R/\pi}\, dt = \frac{\pi M}{\beta}\,(1 - e^{-\beta R}) \le \frac{\pi M}{\beta}. \qquad \square$$

It is natural to define the length of a curve $\gamma(t)$, $t \in [a, b]$, by

$$L(\gamma) = \int_a^b |\gamma'(t)|\, dt$$

(Exercise 3.3.2). This allows us to produce the following estimate:

***3.3.4. Lemma.***

*Let $\gamma$ be a function and $f$ differentiable, defined on $\gamma$, and such that $|f(z)| \le k$ for all $z$. Then*

$$\left|\int_\gamma f(z)\, dz\right| \le k\, L(\gamma).$$

*Proof.* We have

$$\left|\int_\gamma f(z)\, dz\right| = \left|\int_a^b f(\gamma(t))\, \gamma'(t)\, dt\right| \le \int_a^b |f(\gamma(t))|\, |\gamma'(t)|\, dt$$

$$\le k \int_a^b |\gamma'(t)|\, dt = k\, L(\gamma). \qquad \square$$

**3.3.5. Lemma.**

*Suppose that $f$ is holomorphic in $V$ and such that $f'$ is continuous on $V$. Then, for any closed curve $\gamma : [a,b] \to V$,*

$$\oint_\gamma f'(z)\,dz = 0.$$

*In particular,*

$$\oint_\gamma z^n\,dz = 0$$

*if $n \in \mathbb{Z} \setminus \{-1\}$ when $0 \notin \gamma$.*

*Proof.* For each continuous component $\gamma_k$ we have, using the Fundamental Theorem of Calculus,

$$\begin{aligned}\oint_{\gamma_k} f'(z)\,dz &= \int_a^b f'(\gamma_k(t))\,\gamma_k'(t)\,dt \\ &= \int_a^b (f\circ\gamma_k)'(t)\,dt = f(\gamma_k(b)) - f(\gamma_k(a)) = 0.\end{aligned}$$

Then

$$\oint_\gamma f(z)\,dz = \sum_{k=1}^m \oint_{\gamma_k} f(z)\,dz = 0. \qquad \square.$$

The Fundamental Theorem of Calculus usually applies to functions $\mathbb{R} \to \mathbb{R}$. But any proof certainly applies to the case of functions $\mathbb{R} \to \mathbb{C}$, without any changes—just using the absolute value of a complex number instead of the absolute value of a real number.

*3.3.6. Remark*. It is not a mistake that $n = -1$ is excluded in Lemma 3.3.5, and this plays a central role in complex analysis. For now let us notice that if $\gamma(t) = re^{it}$, $t \in [0, 2\pi]$ is a circle around the origin, then

$$\int_\gamma \frac{1}{z}\,dz = \int_0^{2\pi} \frac{ire^{it}}{re^{it}}\,dt = 2\pi i.$$

Things to remark here are the fact that the integral is not zero, that it does not depend on $r$, and that the function $z \mapsto \frac{1}{z}$ has an essential singularity at $z = 0$.

This next result is crucial because, as opposed to Lemma 3.3.5, it does not assume that $f$ has an antiderivative.

***3.3.7. Proposition.***

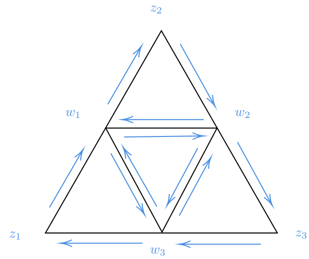
*Let $V \subset \mathbb{C}$ be open, $z_1, z_2, z_3 \subset V$ distinct, and $\gamma = \gamma_{12} \cup \gamma_{23} \cup \gamma_{31}$, where $\gamma_{kj}(t) = tz_j + (1-t)z_k$ is the line segment joining $z_k$ and $z_j$. Suppose that every point in the interior of the triangle is in $V$. Let $z_0 \in V \setminus \gamma$ and $f : V \to \mathbb{C}$ continuous on $V$ and holomorphic on $V \setminus \{z_0\}$. Then*

$$\oint_\gamma f(z)\,dz = 0.$$

*Proof.* In this proof, "triangle" means the area inside the triangle. Assume first that $z_0$ is not in the interior of the triangle $T$ determined by $\gamma$; so for now $f$ is holomorphic everywhere in the interior of the triangle. Consider the middle points

$$w_1 = \frac{z_1 + z_2}{2}, \quad w_2 = \frac{z_2 + z_3}{2}, \quad w_3 = \frac{z_3 + z_1}{2}. \tag{3.5}$$

We can then subdivide the triangle in four triangles.




Let $\ell$ be the perimeter of the original triangle. If we denote by $T_1, T_2, T_3, T_4$ the four triangles thus obtained and $\partial T_k$ the corresponding boundary curves as determined by the arrows, then each $T_k$ has perimeter $\ell/2$ and

$$\oint_\gamma f(z)\,dz = \sum_{k=1}^{4} \oint_{\partial T_k} f(z)\,dz$$

due to the cancellations corresponding to opposing arrows. Let $T_1'$ be the $T_k$ such that $\int_{\partial T_k} f(z)\,dz$ is maximum. We have

$$\left|\oint_\gamma f(z)\,dz\right| \le 4\left|\oint_{\partial T_1'} f(z)\,dz\right|.$$

Now we repeat the process, starting with the triangle $T_1'$. Inductively, we obtain a decreasing sequence $T_1' \supset T_2' \supset \cdots$ of triangles, where $T_k'$ has perimeter $\ell/2^n$, and

$$\left|\oint_\gamma f(z)\,dz\right| \le 4^k\left|\oint_{\partial T_k'} f(z)\,dz\right|. \tag{3.6}$$

Because the sequence $\{T'_k\}$ is decreasing and made out of compact sets, it has nonempty intersection. Let $w_0 \in \bigcap_k T'_k$, and fix $\varepsilon > 0$. As $w_0 \in T$, by hypothesis $f$ is differentiable at $w_0$. So there exists $\delta > 0$ with

$$|f(z) - f(w_0) - f'(w_0)(z - w_0)| \leq \varepsilon\,|z - w_0|, \qquad |z - w_0| < \delta.$$

For $k$ big enough, we have $|z - w_0| < \delta$ for all $z \in T'_k$. Since $z \longmapsto f(w_0) + f'(w_0)(z - w_0)$ is a polynomial, its integral over the perimeter of $T'_k$ is zero by Lemma 3.3.5. Then, using that $|z - w_0| < \ell/2^k$ for all $z \in T'_k$,

$$\begin{aligned}\Big|\oint_{\partial T'_k} f(z)\,dz\Big| &= \Big|\oint_{\partial T'_k} [f(z) - f(w_0) - f'(w_0)(z - w_0)]\,dz \\ &\leq \varepsilon \oint_{\partial T'_k} |z - w_0|\,dz \leq \frac{\varepsilon\ell^2}{4^k}.\end{aligned}$$

Combining with (3.6),

$$\Big|\oint_\gamma f(z)\,dz\Big| \leq \varepsilon\ell^2.$$

As $\varepsilon$ was arbitrary, this shows (3.5). When $z_0$ is not exterior to the triangle, we can subdivide the triangle as a union of smaller triangles in such a way that $z_0$ belongs to a triangle $T_\varepsilon$ as small as we want. By the above the integral is zero is all the other triangles; and since—being continuous on a compact set—$f$ is bounded, say $|f| \leq c$,

$$\Big|\oint_\gamma f(z)\,dz\Big| = \Big|\oint_{T_\varepsilon} f(z)\,dz\Big| \leq c\,m(T_\varepsilon),$$

so $\displaystyle\oint_\gamma f(z)\,dz = 0$. □

The key idea in the proof of Proposition 3.3.7 is the fact, already known from calculus, that differentiable functions can be locally approached by (degree one!) polynomials. It just happens that when differentiation is considered for functions on $\mathbb{C}$, the implications reach further.

Now we can show that every holomorphic function has an antiderivative.

***3.3.8. Proposition.***

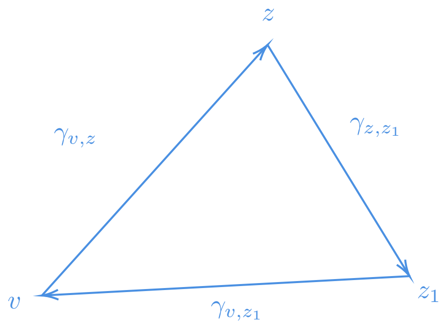
*Let $V \subset \mathbb{C}$ be open and convex, $z_0 \in V$ with $f$ continuous on $V$ and holomorphic on $V \setminus \{z_0\}$. There there exists $F$, holomorphic on $V$, such that $F' = f$ on $V \setminus \{z_0\}$.*

*Proof.* Fix an element $v \in V$. For each $z \in V$, let $\gamma_{v,z}(t) = (1-t)v + tz$, $t \in [0,1]$. As $V$ is convex, each path $\gamma_{v,z}$ is contained in $V$. Define

$$F(z) = \int_{\gamma_{v,z}} f(w)\,dw.$$

Fix $z_1 \in V \setminus \{z_0\}$. Using Proposition 3.3.7 and noting that due to convexity any triangle with vertices in $V$ is contained in $V$,

$$F(z) - F(z_1) = \int_{\gamma_{v,z} \cup (-\gamma_{v,z_1})} f(w)\,dw = -\int_{-\gamma_{z,z_1}} f(w)\,dw = \int_{\gamma_{z_1,z}} f(w)\,dw.$$




Given $\varepsilon > 0$, the continuity of $f$ guarantees that there exists $\delta > 0$ such that $|f(w) - f(z_1)| < \varepsilon$ whenever $|w - z_1| < \delta$. So, for $|w - z_1| < \delta$, since $\int_{\gamma_{z_1,z}} 1\,dw = z - z_1$,

$$\left| \frac{F(z) - F(z_1)}{z - z_1} - f(z_1) \right| = \frac{1}{|z - z_1|} \left| \int_{\gamma_{z_1,z}} [f(w) - f(z_1)]\,dw \right| \leq \varepsilon.$$

Thus the limit of the Newton quotients exists and agrees with $f(z_1)$. That is, $F'(z_1) = f(z_1)$. □

### 3.3.1. Exercises.

(3.3.1) Show that if $\gamma : [a,b] \to \mathbb{C}$ is piecewise continuously differentiable, then it is of bounded variation.

(3.3.2) Let $\gamma(t)$, $t \in [a,b]$ be a curve. Show that it is natural to define the length of $\gamma$ as

$$L(\gamma) = \int_a^b |\gamma'(t)|\,dt$$

## 3.4. The Index

We saw in Remark 3.3.6 that the value of the integral of $1/z$ on a circle around the origin does not depend on the radius of the circle. Playing with this idea will take us far.

**3.4.1. Definition.**

*Let $\gamma : [a, b] \to \mathbb{C}$ be a closed path. For each $z \notin \gamma$, the **index of** $z$ is the number*

$$\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i} \oint_\gamma \frac{1}{w - z}\, dw.$$

*If $\operatorname{Ind}_\gamma(z) \neq 0$, we say that $z$ is **inside** $\gamma$.*

The number $\operatorname{Ind}_\gamma(z)$ is also referred as the **winding number** of $\gamma$ at $z$. We will discuss this name in Remark 3.4.3.

Let us first calculate $\operatorname{Ind}_\gamma(z)$ naively in some examples. These calculations will prove unnecessary very soon, but they are interesting nonetheless. The reader is encouraged to try other approaches—they will likely be cumbersome. The logarithm in this context is not continuous everywhere, so we cannot expect to use the Fundamental Theorem of Calculus.

(1) *A point inside a circle.* Let $\gamma(t) = 5e^{it}$, $t \in [0, 2\pi]$. Let $z = re^{i\theta}$, with $r < 5$. Then

$$\begin{aligned}
\operatorname{Ind}_\gamma(z) &= \frac{1}{2\pi i} \oint_\gamma \frac{1}{w - re^{i\theta}}\, dw = \frac{1}{2\pi i} \int_0^{2\pi} \frac{5ie^{it}}{5e^{it} - re^{i\theta}}\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{1 - \frac{r}{5}\, e^{i(\theta - t)}}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{1 - \frac{r}{5}\, e^{it}}\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} \sum_{k=0}^{\infty} \frac{r^k}{5^k}\, e^{ikt}\, dt = \frac{1}{2\pi} \sum_{k=0}^{\infty} \frac{r^k}{5^k} \int_0^{2\pi} e^{ikt}\, dt \\
&= \frac{1}{2\pi} \int_0^{2\pi} 1\, dt = 1.
\end{aligned}$$

The integral and the series can be exchanged by Dominated Convergence, as the series is bounded in absolute value by constant function $\frac{1}{1 - r/5}$; everything works since $r < 5$.

(2) *A point outside a circle.* Again, let $\gamma(t) = 5e^{it}$, $t \in [0, 2\pi]$. Let $z = re^{i\theta}$, now with $r > 5$. Then—again with the exchange at the end justified

by Dominated Convergence as before—

$$\begin{aligned}\operatorname{Ind}_\gamma(z) &= \frac{1}{2\pi i}\oint_\gamma \frac{1}{w-re^{i\theta}}\,dw = \frac{1}{2\pi i}\int_0^{2\pi}\frac{5ie^{it}}{5e^{it}-re^{i\theta}}\,dt\\ &= -\frac{1}{2\pi}\int_0^{2\pi}\frac{\frac{5}{r}\,e^{i(t-\theta)}}{1-\frac{5}{r}\,e^{i(t-\theta)}}\,dt = -\frac{1}{2\pi}\int_0^{2\pi}\sum_{k=1}^\infty \frac{5^k}{r^k}\,e^{ik(t-\theta)}\,dt\\ &= -\frac{1}{2\pi}\sum_{k=1}^\infty \frac{5^k}{r^k}\int_0^{2\pi} e^{ikt}\,dt = 0.\end{aligned}$$

(3) *A more complicated curve.* Let $\gamma(t) = (1+2\cos t)\,e^{it}$, $t\in[0,2\pi]$, see the plot below. Consider first $z=\frac{3}{2}$. We have (we recommend here to trust a symbolic calculator, rather than suffering the integrals),

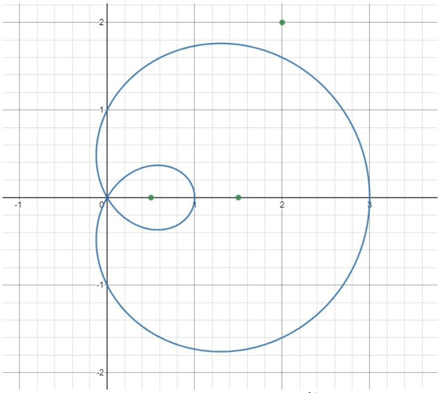
$$\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i}\int_0^{2\pi}\frac{\gamma'(t)}{\gamma(t)-\frac{3}{2}}\,dt = 1.$$

If instead we try $z=\frac{1}{2}$, we obtain

$$\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i}\int_0^{2\pi}\frac{\gamma'(t)}{\gamma(t)-\frac{1}{2}}\,dt = 2.$$

While with $z = 2+2i$, we get

$$\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i}\int_0^{2\pi}\frac{\gamma'(t)}{\gamma(t)-2-2i}\,dt = 0.$$




$\gamma(t) = (1+2\cos t)\,e^{it}$

The fact that in all examples the values obtained were integers was no fluke. Nor the fact that we got zero for those cases where the point was outside the curve. One can also note, in the graph above, that $z=\frac{3}{2}$ is surrounded by the curve once, and that $z=\frac{1}{2}$ is surrounded twice (think about how many times you need to cross the curve to from the point to the region outside the curve); and $\operatorname{Ind}_\gamma(z)$ seems to be accounting for it. After understanding the index map a bit better, we will see in Remark 3.4.3 that this interpretation of the index is intuitively correct.

**3.4.2. Proposition.**

*Let $\gamma$ be a closed path. Then $\operatorname{Ind}_\gamma : \mathbb{C}\setminus\gamma \to \mathbb{C}$ is integer valued, and continuous in each connected component of $\mathbb{C}\setminus\gamma$. There is a single unbounded connected component of $\mathbb{C}\setminus\gamma$, and $\operatorname{Ind}_\gamma = 0$ on that component.*

*Proof.* We have $\gamma = \bigcup_{k=1}^m \gamma_k$ with each $\gamma_k$ closed, continuous, and piecewise differentiable. Since by definition

$$\operatorname{Ind}_\gamma(z) = \sum_{k=1}^m \operatorname{Ind}_{\gamma_k}(z),$$

we may assume without loss of generality that $\gamma$ and $\gamma'$ are continuous.

Fix $z \in \mathbb{C}\setminus\gamma$. By definition of the integral,

$$\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)-z}\,dt.$$

Define $g : [a,b] \to \mathbb{C}$ by

$$g(t) = \exp\left[\int_a^t \frac{\gamma'(s)}{\gamma(s)-z}\,ds\right].$$

As the integrand is continuous, we have that $g$ is continuous, and that

$$g'(t) = \frac{g(t)\,\gamma'(t)}{\gamma(t)-z}$$

for $t \in [a,b]$. We may rewrite this as

$$0 = g'(t)(\gamma(t)-z) - g(t)\gamma'(t) = (\gamma(t)-z)^2\left(\frac{g(t)}{\gamma(t)-z}\right)'.$$

Since $\gamma(t) - z \neq 0$ for all $t$ by hypothesis, the function $\frac{g(t)}{\gamma(t)-z}$ is continuous on $[a,b]$ and its derivative is zero; thus it is constant. As $g(a) = 1$, we have

$$\frac{g(t)}{\gamma(t)-z} = \frac{1}{\gamma(a)-z}.$$

Thus

$$g(t) = \frac{\gamma(t)-z}{\gamma(a)-z}.$$

As $\gamma$ is closed, $\gamma(b) = \gamma(a)$ and so $g(b) = 1$. This equality is

$$\exp(2\pi i\ \operatorname{Ind}_\gamma(z)) = 1,$$

which can only occur if $\operatorname{Ind}_\gamma(z) \in \mathbb{Z}$.

As $\gamma$ is compact, since $z \in \mathbb{C}\setminus\gamma$ there is a small disk $z + r\,\mathbb{D}$ around $z$ such that $\operatorname{dist}(z+r\,\mathbb{D}, \gamma) > \delta$ for some $\delta > 0$. Then the function $w \longmapsto \frac{1}{\gamma(t)-w}$

is uniformly continuous on $z + r\,\mathbb{D}$, independently of $t$. It follows from this, together with the fact that $\gamma'$ is bounded (being continuous on a compact set), that $\mathrm{Ind}_\gamma(z)$ is continuous on the connected components of $\mathbb{C} \setminus \gamma$. As a continuous function maps connected sets to connected sets, we deduce that $\mathrm{Ind}_\gamma(z)$, being integer-valued, is constant on connected components.

Since $\gamma$ is compact, it is contained in some disk $r\,\mathbb{D}$. Then $\mathbb{C} \setminus r\,\mathbb{D}$ is contained in a single connected component of $\mathbb{C} \setminus \gamma$. That is, there is a single unbounded connected component of $\mathbb{C} \setminus \gamma$, and it is clear that $\mathrm{Ind}_\gamma(z) = 0$ on that component since it is constant and we can consider $z$ with arbitrarily large $|z|$. $\square$

*3.4.3. Remark*. The name *winding number* for $\mathrm{Ind}_\gamma(z)$ can be understood as follows. By definition, if $\gamma : [a, b] \to \mathbb{C}$ is a closed curve,

$$\int_a^b \frac{\gamma'(t)}{\gamma(t) - z}\, dt = 2\pi i \ \mathrm{Ind}_\gamma(z).$$

Assume we can write $\gamma(t) = z + r(t)\, e^{i\theta(t)}$ for differentiable real functions $r$ and $\theta$ (the continuity of $\theta$ might not be that easy to establish, but this is an intuitive argument and not a proof). Then $\gamma'(t) = r'(t)\, e^{i\theta(t)} + i\, r(t)\, \theta'(t)\, e^{i\theta(t)}$. Thus, as $r(b) = r(a)$,

$$\begin{aligned} 2\pi i \ \mathrm{Ind}_\gamma(z) &= \int_a^b \frac{(r'(t) + i\, r(t)\, \theta'(t))\, e^{i\theta(t)}}{r(t)\, e^{i\theta(t)}}\, dt \\ &= \int_a^b \frac{r'(t)}{r(t)}\, dt + i \int_a^b \theta'(t)\, dt \\ &= i \int_a^b \theta'(t)\, dt = i\, [\theta(b) - \theta(a)] \end{aligned}$$

So

$$\mathrm{Ind}_\gamma(z) = \frac{\theta(b) - \theta(a)}{2\pi} \tag{3.7}$$

counts the number of times that $\theta(t)$ passes by $2\pi$, that is the "number of full turns" of the curve around $z$. This assumes that the curve turns in counter-clockwise direction (i.e., $\theta$ increases). When the curve turns the other way, the winding number will reflect this by being negative.

In the particular case where $\gamma(t) = z_0 + re^{it}$, $t \in [0, 2\pi]$ is a circle,

$$\mathrm{Ind}_\gamma(z) = \begin{cases} 1, & |z - z_0| < r, \\ 0, & |z - z_0| > r \end{cases}$$

Indeed, by Proposition 3.4.2 we have that $\mathrm{Ind}_\gamma(z) = \mathrm{Ind}_\gamma(z_0)$ when $|z - z_0| < r$, and $\mathrm{Ind}_\gamma(z) = 0$ when $|z - z_0| > r$ (i.e., outside the circle). Then

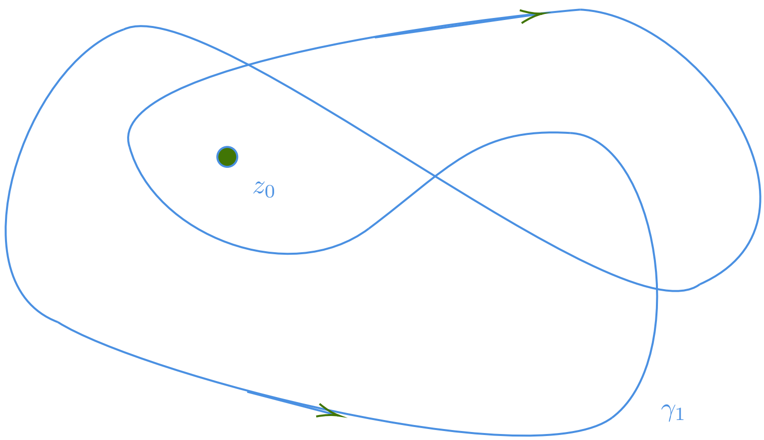
$$\mathrm{Ind}_\gamma(z) = \mathrm{Ind}_\gamma(z_0) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{ir\, e^{it}}{re^{it}}\, dt = 1$$

Calculating the index by brute force for most curves looks like a practical impossibility. It is also often not possible to use (3.7) because there is no guarantee that the argument of a continuous curve $\gamma(t) = r(t)\, e^{i\theta(t)}$ is continuous. But most of the times we will only need it over circles, so the problems do not arise.

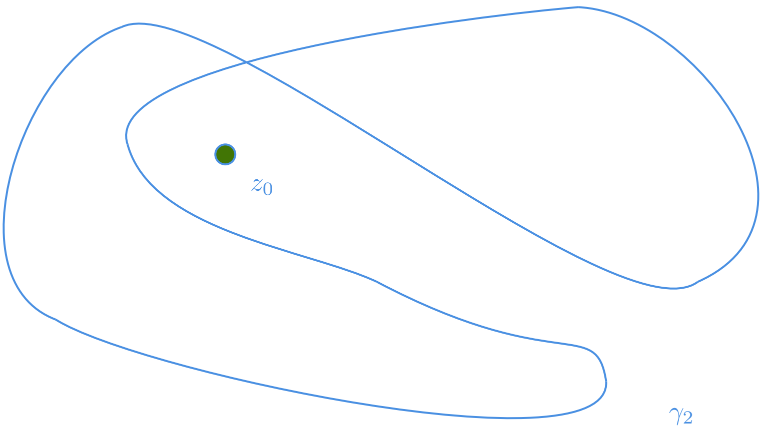
It is possible to show that if two curves $\gamma_1$ and $\gamma_2$ have common endpoints and $\gamma_1$ can be changed continuously into $\gamma_2$ (that is, they are **homotopic**) then

$$\int_{\gamma_1} f(z)\, dz = \int_{\gamma_2} f(z)\, dz \tag{3.8}$$

for any function $f$ holomorphic on an open set that contains both curves. Then for instance if we ask ourselves what $\mathrm{Ind}_{\gamma_1}(z_0)$ is in the following situation:




it is not hard to do it intuitively, if one pays attention to the changes of direction; but it can be confusing. But (3.8) guarantees that $\mathrm{Ind}_{\gamma_1}(z_0) = \mathrm{Ind}_{\gamma_2}(z_0)$, where




and now it is obvious that $\operatorname{Ind}_{\gamma_2}(z_0) = 0$.

## 3.5. Cauchy's Theorem

We have now done all the heavy lifting to prove what is arguably the most crucial theorem in complex analysis. The convexity hypothesis is only a crutch that will be dropped soon.

***3.5.1. Theorem. (Cauchy's Theorem and Integral Formula, First Version.)***

*Let $V \subset \mathbb{C}$ be open and convex, $\gamma$ a closed curve in $V$, and $f$ holomorphic on $V$. Then*

$$\oint_\gamma f(z)\, dz = 0, \tag{3.9}$$

*and*

$$f(z) \operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(w)}{w - z}\, dw, \qquad z \in V \setminus \gamma. \tag{3.10}$$

*Proof.* The equality (3.9) follows from a direct application of Proposition 3.3.8 and Lemma 3.3.5. For the second equality, given $z \in V$ define

$$g(w) = \begin{cases} \frac{f(w)-f(z)}{w-z}, & w \in V \setminus \{z\} \\ f'(z), & w = z \end{cases}$$

Then $g$ is continuous on $V$, and holomorphic on $V \setminus \{z\}$; by (3.9),

$$\oint_\gamma g_z(w)\, dw = 0.$$

Thus

$$\oint_\gamma \frac{f(w)}{w-z}\, dw = \oint_\gamma \frac{f(z)}{w-z}\, dw = 2\pi i\, f(z)\, \mathrm{Ind}_\gamma(z). \qquad \square$$

Theorem 3.5.1 will often be enough generality, as typically $V$ is a disk and $\gamma$ a circle. But as we mentioned, convexity can be dropped, as it still happens locally (cfr. Theorem 3.9.1). And here is the first big application of Cauchy's Integral Formula:

**3.5.2. Corollary.**

*Let $V \subset \mathbb{C}$ be open and $f : V \to \mathbb{C}$. Then $f$ is holomorphic if and only if it is analytic. For any $z_0 \in V$, there exist coefficients $\{a_k\}$ such that $f(z) = \sum_{k=0}^\infty a_k (z - z_0)^k$ in any disk $z_0 + r\mathbb{D}$ contained in $V$.*

*Proof.* We already know from Proposition 3.1.3 that if $f$ is analytic, then it is holomorphic. So assume that $f$ is holomorphic. Fix $z_0 \in V$; as $V$ is open, we can choose $r > 0$ such that $z_0 + r\,\mathbb{D} \subset V$. This $r$ doesn't have to small. On the disk $z_0 + r\,\mathbb{D}$ we can apply Theorem 3.5.1. So if $\gamma(t) = z_0 + s\, e^{it}$, $t \in [0, 2\pi]$ and $s < r$, then

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(w)}{w-z}\, dw, \qquad z \in z_0 + s\,\mathbb{D}.$$

Now, noting that $|z - z_0| < s$ and that the uniform convergence of the series guarantees the exchange with the integral,

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w-z}\, dw = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(\gamma(t))\, ie^{it}}{z_0 - z + se^{it}}\, dt \\ &= \frac{1}{2\pi s} \int_0^{2\pi} \frac{f(\gamma(t))}{1 - \frac{z-z_0}{s}\, e^{-it}}\, dt = \frac{1}{2\pi s} \int_0^{2\pi} f(\gamma(t)) \sum_{k=0}^\infty \big(\frac{z-z_0}{s}\, e^{-it}\big)^k\, dt \\ &= \sum_{k=0}^\infty a_k\, (z - z_0)^k, \end{aligned}$$

where
$$a_k = \frac{1}{2\pi s^{k+1}} \int_0^{2\pi} f(z_0 + se^{it})\, e^{-ikt}\, dt, \qquad k \in \mathbb{N} \cup \{0\}.$$
The coefficients $a_k$ do not depend on $z$. Proposition 3.1.3 guarantees that they are unique and that they do not depend on $s$, so the power series representation works in all of $z_0 + r\,\mathbb{D}$. □

Once we know that a holomorphic $f$ is analytic as in Corollary 3.5.2, it is immediate that
$$a_k = \frac{f^{(k)}(z_0)}{k!}, \qquad k \in \{0\} \cup \mathbb{N},$$
in any disk centered at $z_0$ and contained in $V$, showing again that the coefficients do not depend on $s$.

The converse of Cauchy's Theorem actually holds, and is usually considered as a result on its own.

***3.5.3. Corollary. (Morera's Theorem)***

*Let $V \subset \mathbb{C}$ be open and $f : V \to \mathbb{C}$ such that*
$$\oint_\gamma f(z)\, dz = 0$$
*for every triangle $\gamma \subset V$. Then $f$ is holomorphic on $V$.*

*Proof.* Since being holomorphic at a point is a local property and $V$ is open, once we fix $z \in V$ we may assume that $V$ is a disk; so convex. Now we can reproduce the proof of Proposition 3.3.8, by using the hypothesis in place of Proposition 3.3.7, to show that $f$ has an antiderivative. This antiderivative is then holomorphic by definition, and analytic by Corollary 3.5.2; thus its derivative $f$ is analytic, and thus holomorphic. At this stage the disk matters no more, as we have shown that $f'(z)$ exists for all $z \in V$. □

## 3.6. Zeros of Holomorphic Functions

Finding and understanding roots (or zeros) of polynomials has been one of the first motivations of algebra and a pillar of its developments, from Cardano's formula to Galois Theory, Hilbert's Nullstellensatz, and beyond. Here it makes sense for us to pay attention to zeros of holomorphic functions.

**3.6.1. Definition.**

*Let $V \subset \mathbb{C}$ be open and $f$ holomorphic on $V$. A **zero** for $f$ is any $w \in V$ such that $f(w) = 0$. If $f(z) = (z-w)^m g(z)$ with $g$ holomorphic and $g(w) \neq 0$, we say that the **order** of $w$ is $m$. When we list the zeros of $f$ **counting multiplicities** we are repeating each zero as many times as its order.*

For an easy example of what *counting multiplicities* means, if we list the zeros of $f(z) = z^2(z-1)(z-2)^3$ counting multiplicities, the list would be $0, 0, 1, 2, 2, 2$. The order is irrelevant, what matters is the number of times each zero appears in the list.

The order is well-defined, in the sense that if $(z-w)^n g(z) = (z-w)^m h(z)$ with $g, h$ holomorphic at $w$ and $g(w)h(w) \neq 0$, then $n = m$ (Exercise 3.6.1). It remains to be seen that every zero of a holomorphic function has well-defined order. We will answer this very soon.

***3.6.2. Corollary.***

*Let $V \subset \mathbb{C}$ be open and connected, and $f$ holomorphic on $V$. Then the following statements are equivalent:*

*(i)* $f = 0$;

*(ii) there exists* $z_0 \in V$ *such that* $f^{(n)}(z_0) = 0$ *for all* $n \in \mathbb{N} \cup \{0\}$;

*(iii) the set* $\{z \in V : \ f(z) = 0\}$ *has a cluster point in* $V$.

*Proof.* (i) $\implies$ (iii): this is trivial.

(iii) $\implies$ (ii): let $z_0$ be a cluster point of $\{z \in V : \ f(z) = 0\}$. As $f$ is continuous, $f(z_0) = 0$. Suppose that there exists $n$ with $f^{(n)}(z_0) \neq 0$; we may assume that $n$ is the minimum such $n$, that is $f^{(k)}(z_0) = 0$ for $k = 0, \ldots, n-1$ and $f^{(n)}(z_0) \neq 0$. By Corollary 3.5.2 we know that $f$ is analytic at $z_0$, and that on some disk around $z_0$,

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)\,(z-z_0)^k}{k!} = \sum_{k=n}^{\infty} \frac{f^{(k)}(z_0)\,(z-z_0)^k}{k!}.$$

So we can write $f(z) = (z-z_0)^n\, g(z)$, with $g$ analytic and $g(z_0) = a_n \neq 0$. By continuity, there exists $r > 0$ such that $g(z) \neq 0$ for all $z \in z_0 + r\,\mathbb{D}$. As $z_0$ is a cluster point of the zeros of $f$, there exists $z_1 \in z_0 + r\,\mathbb{D}$, $z_1 \neq z_0$, with $f(z_1) = 0$. Then $0 = (z_1 - z_0)g(z_1)$, giving us $g(z_1) = 0$, a contradiction. It follows that $f^{(k)}(z_0) = 0$ for all $k$.

(ii) $\implies$ (i): let

$$W = \{z \in V : \ f^{(k)}(z) = 0, \ k \in \mathbb{N} \cup \{0\}\}.$$

By hypothesis, $W \neq \varnothing$. Also, $W$ is closed, since $f^{(k)}$ is continuous for all $k$. If $z_1 \in W$, since $f$ is analytic at $z_1$ there exists $r > 0$ such that

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_1)\,(z - z_1)^k}{k!} = 0, \qquad z \in z_1 + r\,\mathbb{D}.$$

So $z_1 + r\,\mathbb{D} \subset W$, and $W$ is open. Being clopen in the connected set $V$, necessarily $W = V$. ☐

The connectivity hypothesis in Corollary 3.6.2 is not just an artifact of the proof. If $V = \mathbb{D} \cup (2 + \mathbb{D})$, say, then the function $f = 1_{\mathbb{D}}$ is holomorphic on $V$, it satisfies the conditions (ii) and (iii), but it does not satisfy condition (i).

One consequence of Corollary 3.6.2 is that if $V$ is bounded and a holomorphic function $f$ has infinitely many zeros in $V$, then these zeros have to converge towards the boundary as they can have no cluster point inside $V$.

***3.6.3. Corollary.***

*Let $V \subset \mathbb{C}$ be open and connected, and $f$ holomorphic on $V$ and nonzero. Then the zeros of $f$ are isolated, and they have finite multiplicity in the sense that if $f(z_0) = 0$, then there exists $n \in \mathbb{N}$ such that $f(z) = (z - z_0)^n g(z)$ with $g$ holomorphic and $g(z_0) \neq 0$.*

*In particular, a nonzero holomorphic function has finitely many zeros in any closed disk contained in its domain.*

*Proof.* By Corollary 3.6.2, the set of zeros of $f$ has no cluster points. If $f(z_0) = 0$, again by Corollary 3.6.2 there exists $n \in \mathbb{N}$ such that $f^{(n)}(z_0) \neq 0$. Then, as in the proof of Corollary 3.6.2, we can write $f(z) = (z - z_0)^n g(z)$, with $g(z_0) \neq 0$. ☐

***3.6.4. Corollary.***

*Let $V \subset \mathbb{C}$ be open and connected, and $f$ a nonzero holomorphic function on $V$. Let $z_1, \ldots, z_m$ be the zeros of $f$ inside some closed disk $D$, and $n_1, \ldots, n_m$ their orders. Then there exists $g$ holomorphic and nonzero on $D$, and such that*

$$f(z) = (z - z_1)^{n_1} \cdots (z - z_m)^{n_m} g(z), \qquad z \in D.$$

*Proof.* By Corollary 3.6.3, we have $f(z) = (z - z_1)^{n_1} g_1(z)$, where $n_1$ is the multiplicity of $z_1$ and $g$ is holomorphic with $g_1(z_1) \neq 0$. If $z_{n_1+1}$ is the next zero (i.e., if $z_2 = \cdots = z_{n_1} = z_1$), we get $g_1(z) = (z - z_{n_1+1})^{n_2} g_2(z)$, where $g_2$ is holomorphic and $g_2(z_{n_1+1}) \neq 0$ (and $g_2(z_1) \neq 0$, also). Repeating this, we may write $f(z) = (z - z_1)^{n_1} \cdots (z - z_m)^{n_m} g(z)$, where $g = g_m$ is holomorphic and nonzero on $D$. □

It is worth mentioning that we are still requiring convexity in some results. This is a crutch that will be dropped on Section 3.9.

**3.6.5. Corollary.**

*Let $V \subset \mathbb{C}$ be open and convex, and $f$ a holomorphic function on $V$ with zeros $z_1, \ldots, z_m$ (counting multiplicities). If $\gamma$ is a closed curve on $V$ with $\{z_1, \ldots, z_m\} \cap \gamma = \varnothing$, then*

$$\oint_\gamma \frac{f'(z)}{f(z)}\, dz = 2\pi i \sum_{k=1}^{m} \operatorname{Ind}_\gamma(z_k).$$

*Proof.* Note that $V$, being convex, is path connected and hence connected. By Corollary 3.6.4 we may write, counting multiplicities,

$$f(z) = (z - z_1) \cdots (z - z_m) g(z), \tag{3.11}$$

where $g$ is holomorphic and $g(z_k) \neq 0$ for all $k$. Differentiating (3.11),

$$\begin{aligned} f'(z) &= (z - z_1) \cdots (z - z_m) g'(z) + \sum_{k=1}^{m} \frac{1}{z - z_k} \prod_{j=1}^{m} (z - z_k)\, g(z) \\ &= f(z) \left[ \frac{g'(z)}{g(z)} + \sum_{k=1}^{m} \frac{1}{z - z_j} \right]. \end{aligned}$$

As $\frac{g'(z)}{g(z)}$ is holomorphic on $V$, by (both parts of) Theorem 3.5.1 we have

$$\int_\gamma \frac{f'(z)}{f(z)}\, dz = \int_\gamma \frac{g'(z)}{g(z)}\, dz + \sum_{k=1}^{m} \int_\gamma \frac{1}{z - z_k}\, dz = 2\pi i \sum_{k=1}^{m} \operatorname{Ind}_\gamma(z_k).$$ □

It is not hard to find functions which are holomorphic in $\mathbb{D}$ and have infinitely many zeros. For instance consider

$$f(z) = \sin \frac{1}{1 - z}, \qquad z \in \mathbb{D},$$

with zeros $z_k = 1 - \frac{1}{2k\pi}$, $k \in \mathbb{N}$. Taking $w_n = 1 - \frac{i}{n}$ shows that

$$\sin \frac{1}{1 - w_n} = \sin(-in) = \frac{e^n - e^{-n}}{2i}$$

is unbounded.

### 3.6.1. Exercises.

(3.6.1) Show that the order of a zero is well-defined, in the sense that if $(z-w)^n g(z) = (z-w)^m h(z)$ with $g, h$ holomorphic at $w$ and $g(w)h(w) \neq 0$, then $n = m$.

(3.6.2) Let $V$ be open and $f$ holomorphic on $V$. Show that $f$ has at most countably many zeros.

(3.6.3) Let $V$ be a simply connected region and $f$ holomorphic on $V$ such that it has no zeros on $V$. Show if $a \in V$ and

$$g(z) = \log f(a) + \int_{\gamma_{a,z}} \frac{f'(w)}{f(w)}\, dw,$$

then $f(z) = e^{g(z)}$. Note that we know from the proof of Proposition 3.3.8 how to differentiate the integral.

## 3.7. Maximum Modulus Principle and Liouville's Theorem

We keep taking advantage of the dual nature of analytic/holomorphic functions.

***3.7.1. Proposition.***

*If*

$$f(z) = \sum_k a_k (z-z_0)^k, \qquad z \in z_0 + r\,\mathbb{D},$$

*with radius of convergence $r > 0$, then for any $0 < s < r$*

$$\sum_{k=0}^{\infty} |a_k|^2\, s^{2k} = \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + se^{it})|^2\, dt.$$

*Proof.* We have the equality

$$f(z_0 + se^{it}) = \sum_k a_k s^k e^{ikt}.$$

The series converges absolutely and uniformly by Proposition 3.1.3, since $0 < s < r$. Then by Dominated Convergence, as the map $t \longmapsto f(z_0 + se^{it})$ is continuous and hence bounded,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + se^{it})|^2\, dt = \frac{1}{2\pi} \int_0^{2\pi} \Big| \sum_k a_k s^k\, e^{ikt} \Big|^2\, dt$$

$$
\begin{aligned}
&= \frac{1}{2\pi} \int_0^{2\pi} \sum_{k,j} a_k \, \overline{a_j} \, s^{2k} \, e^{i(k-j)t} \, dt \\
&= \frac{1}{2\pi} \sum_{k,j} a_k \, \overline{a_j} \, s^{2k} \int_0^{2\pi} e^{i(k-j)t} \, dt \\
&= \frac{1}{2\pi} \sum_k |a_k|^2 \, s^{2k} \, 2\pi \\
&= \sum_k |a_k|^2 \, s^{2k}. \qquad \square
\end{aligned}
$$

As have already seen and we will keep seeing time and time again, Cauchy's Integral Formula has far reaching implications. Here is one that would be hard to imagine by just looking at Definition 3.1.2.

***3.7.2. Corollary. (Maximum Modulus Principle)***

*If $V$ is open and $f$ is holomorphic on $V$ and non-constant, the value $\sup\{|f(z)| : z \in V\}$ is never attained. In other words, if $f$ attains its maximum in $V$, then $f$ is constant.*

*Proof.* Suppose that $|f(z_0)| \geq |f(z)|$ for all $z \in V$. Choose $r > 0$ such that $z_0 + r\,\mathbb{D} \subset V$. Using Cauchy's Formula (3.10) on the boundary of the disk we have

$$
\begin{aligned}
|f(z_0)| &= \left| \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{it})}{re^{it}} \, ir \, e^{it} \, dt \right| = \frac{1}{2\pi} \left| \int_0^{2\pi} f(z_0 + re^{it}) \, dt \right| \\
&\leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{it})| \, dt \leq |f(z_0)|.
\end{aligned}
$$

The last inequality is due to the assumption that $|f(z)| \leq |f(z_0)|$ in $V$. From this we deduce

$$
\int_0^{2\pi} \big[\, |f(z_0)| - |f(z_0 + re^{it})| \,\big] \, dt = 0
$$

and the integrand is non-negative by hypothesis; so $|f(z_0 + re^{it})| = |f(z_0)|$ for all $t$. The proof of Corollary 3.5.2 guarantees that $f(z) = \sum_k a_k (z - z_0)^k$ on any disk $z_0 + r\mathbb{D}$ as long as the disk is contained in $V$. By Proposition 3.7.1,

$$
\sum_k |a_k|^2 \, r^{2k} = \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{it})|^2 \, dt = |f(z_0)|^2 = |a_0|^2.
$$

Then $a_k = 0$ for all $k \geq 1$, implying that $f$ is constant in $z_0 + r\,\mathbb{D}$. Thus the analytic function $f$ agrees with the constant function $f(z_0)$ on a set that contains an accumulation point; by Corollary 3.6.2, $f(z) = f(z_0)$ for all $z \in V$. □

We can phrase Corollary 3.7.2 as saying that a holomorphic function never attains a local maximum or minimum (in absolute value) inside a disk, unless it is constant.

A look at the proof of Corollary 3.7.2 shows that we proved the following:

**3.7.3. Corollary.**

*Let $f$ be holomorphic in an open set $V$, $z_0 \in V$ and $r > 0$ such that $z_0 + r\mathbb{D} \subset V$. Then, with $\gamma$ the circle of radius $r$ around $z_0$,*

$$|f(z_0)| < \max\{|f(w)| :\ w \in \gamma\}. \tag{3.12}$$

The requirement that $f$ is holomorphic in the whole interior of the circle in Corollary 3.7.3 is not superfluous. Let $f(z) = \frac{1}{z}$, $\gamma$ the circle of radius 2 around 0, and $z_0 = 1$. Then $f$ is holomorphic on the annulus $\frac{1}{2} < |z| < 3$ that contains $\gamma$ and $z_0$, but $|f(z_0)| = 1$ while $|f(w)| = \frac{1}{2}$ for all $w \in \gamma$.

Here is a very useful consequence of Proposition 3.7.1. The result is most likely known to the reader in the case of polynomials. A function that is holomorphic in all of $\mathbb{C}$ is called **entire**.

**3.7.4. Corollary. (Liouville's Theorem)**

*Let $f$ be holomorphic on $\mathbb{C}$. If $f$ is bounded, then $f$ is constant.*

*Proof.* Suppose $|f(z)| \leq c$ for all $z \in \mathbb{C}$. By the proof of Corollary 3.5.2,

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \qquad z \in \mathbb{C},$$

and by Proposition 3.7.1 for any $r > 0$

$$\sum_{k=0}^{\infty} |a_k|^2 r^{2k} = \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^2 \, dt \leq c^2.$$

We get $c^2 \geq |a_k|^2 r^{2k}$ for all $r > 0$, which implies $a_k = 0$. As this works for all $k \geq 1$, $a_1 = a_2 = \cdots = 0$, and so $f$ is constant. □

**3.7.5. Corollary. (Fundamental Theorem of Algebra)**

*Let $p \in \mathbb{C}[x]$ be a polynomial with $\deg p \geq 1$. Then there exists $z_0 \in \mathbb{C}$ with $p(z_0) = 0$. As a consequence, any complex polynomial $p$ of degree $n$*

*can be written as*

$$p(z) = a\,(z - z_1)\cdots(z - z_n).$$

*Proof.* Suppose that $p(z) \neq 0$ for all $z \in \mathbb{C}$. We lose no generality if we assume that $p$ is monic. So $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$. Let $g(z) = 1/p(z)$ and $c = \max\{|a_0|, \ldots, |a_{n-1}|\}$. Then

$$|p(re^{it})| = \Big|r^n e^{int} + \sum_{k=0}^{n-1} a_k r^k e^{ikt}\Big| \geq r^n - \sum_{k=0}^{n-1} c\,r^k = \frac{r^{n+1} - (1+c)r^n + c}{r-1} > 0 \tag{3.13}$$

if $r$ is big enough. Say $r_0 > 1$ is one such $r$. The function

$$h(r) = \frac{r-1}{r^{n+1} - (1+c)r^n + c}, \qquad r > r_0,$$

is positive and decreases to 0 as $r \to \infty$. This implies that there exists $c' > 0$ with $h(r) < c'$ for all $r \geq r_0$. Thus (3.13) gives us

$$|g(z)| \leq h(|z|) < c', \qquad |z| \geq r_0.$$

So $g$ is bounded both inside the disk $\overline{r_0\,\mathbb{D}}$—being a continuous function on a compact set—and on its complement. As an entire bounded function, $g$ is constant by Liouville's Theorem and so $p$ is constant, contradicting that $\deg p \geq 1$. Thus there exists $z_1$ such that $p(z_1) = 0$. Using Corollary 3.6.3 (or simply dividing $p$ by $z - z_1$ enough times) we get that $p(z) = (z - z_1)^{n_1}\, g(z)$, with $n_1 \in \mathbb{N}$ and $g$ a polynomial with $\deg g < \deg p$ and $g(z_1) \neq 0$. We can now apply the argument above to $g(z)$, to obtain a root $z_2$. Continuing inductively, as the degree decreases every time,

$$p(z) = (z - z_0)^{n_1} \cdots (z - z_r)^{n_r}.$$

□

*3.7.6. Remark*. It is tempting to try and prove Liouville's Theorem by using the Maximum Modulus Principle. But a direct proof is not possible. Consider the function $f : \mathbb{C} \to \mathbb{C}$ given by

$$f(z) = \frac{z}{1 + |z|}.$$

For any $z_0 \in \mathbb{C}$, write $z_0 = re^{i\theta}$. For any $\delta > 0$, if $0 < s < \delta$ we have

$$|f(z_0)| = \frac{r}{1+r} < \frac{r+s}{1+|r+s|} = \frac{|z_0 + se^{i\theta}|}{1 + |z_0 + se^{i\theta}|} = |f(z_0 + se^{i\theta})|.$$

We have used that the real function $t \longmapsto \frac{t}{1+t}$ is strictly increasing for $t > 0$. The inequality above shows that $f$ has no local maximum, in absolute value, in any disk that contains $z_0$; and the same idea works for minima if we use $-s$ instead of $s$. Thus, $f$ is not holomorphic but it satisfies the conclusion of the Maximum Modulus Principle, while being bounded and not-constant. What this shows is that a proof of Liouville's Theorem needs to use that $f$ is holomorphic in a deeper way than just satisfying the Maximum Modulus Principle.

Let $\mathbb{A}_B$ be the algebra of holomorphic bounded functions on $\mathbb{D}$. Equipped with the supremum norm, $\mathbb{A}_B$ becomes a complete algebra (a Banach algebra, in terminology to be defined in Section 9.2). To check for completeness we need to see that a uniform limit of bounded holomorphic functions is holomorphic. So suppose that $\{f_n\}$ is Cauchy in $\|\cdot\|_\infty$. Let $\gamma$ be any closed curve in $\mathbb{D}$. By the uniform convergence,

$$\oint_\gamma f(z)\,dz = \int_\gamma \lim_n f_n(z)\,dz = \lim_n \int_\gamma f_n(z)\,dz = 0.$$

Then Morera's Theorem (Corollary 3.5.3) implies that the function $f$ is analytic. We have shown

***3.7.7. Proposition.***

*Let $V \subset \mathbb{C}$ be open and $\{f_n\}$ a sequence of bounded holomorphic functions on $V$. If $f_n \to f$ uniformly on $V$, then $f$ is holomorphic on $V$.*

This last result emphasizes one big different between complex and real analytic functions. Indeed, on an interval $[a, b] \subset \mathbb{R}$, any continuous function is a uniform limit of polynomials (see Theorem 7.4.20 and Remark 7.4.21).

## 3.8. Consequences of Cauchy's Theorem

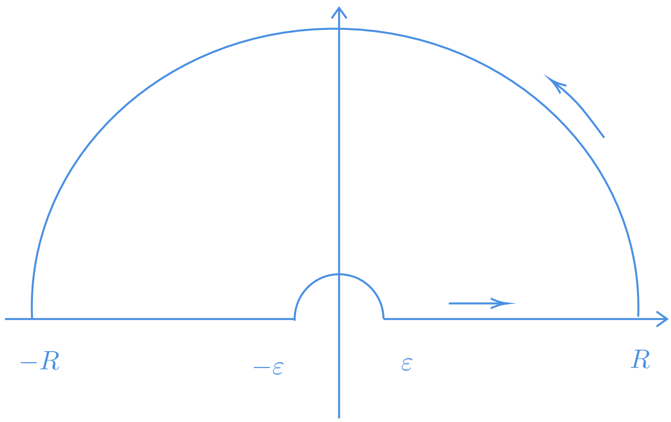
The current version we have of Cauchy's Theorem has proven very useful, even though it has the restriction that it requires the domain $V$ to be convex besides being open. We will prove the general version soon, but let us first obtain some more mileage out of the initial version.

*3.8.1. Example* (*A well-known integral*). We want to calculate $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$. By definition of the improper integral and the sine,

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx = \lim_{R\to\infty} \lim_{\varepsilon\to 0} \int_{-R}^{-\varepsilon} \frac{\sin x}{x}\,dx + \int_{\varepsilon}^{R} \frac{\sin x}{x}\,dx$$

$$= \lim_{R\to\infty} \lim_{\varepsilon\to 0} \operatorname{Im} \int_{-R}^{-\varepsilon} \frac{e^{ix}}{x}\,dx + \int_{\varepsilon}^{R} \frac{e^{ix}}{x}\,dx.$$

Fix $R > \varepsilon > 0$. Consider the curve $\gamma(t)$




Note that $f(z) = \frac{e^{iz}}{z}$ is holomorphic in a neighbourhood of $\gamma$. By Cauchy's Theorem 3.5.1, $\int_\gamma f(z)\,dz = 0$. We can consider $\gamma$ as the join of four curves, two segments and two semicircles. On $[-R, -\varepsilon]$, we can parametrize as $\gamma_1(t) = t$, $t \in [-R, -\varepsilon]$. Working similarly on $\gamma_3(t) = t$ for $t \in [\varepsilon, R]$,

$$\oint_{\gamma_1} \frac{e^{iz}}{z}\,dz = \int_{-R}^{-\varepsilon} \frac{e^{ix}}{x}\,dx, \qquad \int_{\gamma_3} \frac{e^{iz}}{z}\,dz = \int_{\varepsilon}^{R} \frac{e^{ix}}{x}\,dx.$$

On the small semicircle, $\gamma_2(t) = \varepsilon\, e^{i(\pi - t)}$, $t \in [0, \pi]$. Using Dominated Convergence,

$$\int_{\gamma_2} \frac{e^{iz}}{z}\,dz = -\int_0^\pi e^{i\varepsilon\cos(\pi - t)}\, e^{-\varepsilon\sin(\pi - t)}\, i\,dt \xrightarrow[\varepsilon\to 0]{} -\pi i.$$

On the big semicircle, with $\gamma_4(t) = R\,e^{it}$, $t \in [0, \pi]$, we can apply Jordan's Lemma 3.3.3 to get

$$\left|\int_{\gamma_4} \frac{e^{iz}}{z}\,dz\right| \le \frac{\pi}{R},$$

showing that this integral goes to zero as $R \to \infty$. Taking limit on $\varepsilon$ and $R$ in the equality

$$0 = \sum_{k=1}^{4} \int_{\gamma_k} \frac{e^{iz}}{z}\,dz,$$

we have shown that

$$\int_{-R}^{-\varepsilon} \frac{e^{ix}}{x}\,dx + \int_{\varepsilon}^{R} \frac{e^{ix}}{x}\,dx \xrightarrow[R\to\infty,\varepsilon\to 0]{} i\pi.$$

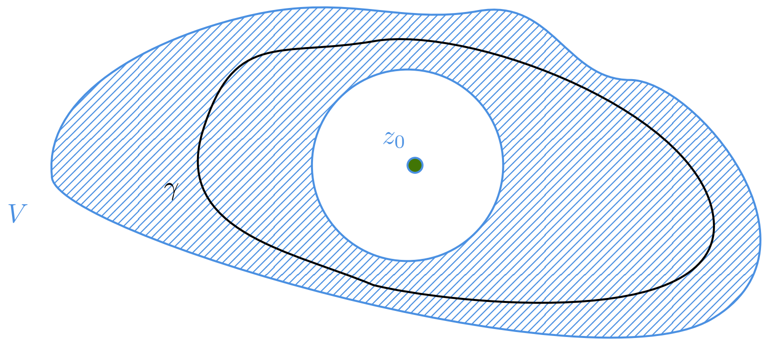
Then, as $\frac{\sin x}{x}$ is an even function,

$$\int_0^\infty \frac{\sin x}{x}\,dx = \tfrac{1}{2}\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dt = \tfrac{1}{2}\lim_{R\to\infty}\lim_{\varepsilon\to 0} \operatorname{Im}\left(\int_{-R}^{-\varepsilon} \frac{e^{ix}}{x}\,dx + \int_{\varepsilon}^{R} \frac{e^{ix}}{x}\,dx\right) = \frac{\pi}{2}.$$

## 3.9. The General Cauchy Theorem

In Theorem 3.5.1 we required that $V$ be convex. It is not the time to replace that condition with a more relaxed one, that takes the curve $\gamma$ into account.

Given a closed curve $\gamma$ in an open set $V$, the new condition will be that $\operatorname{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$. This could mean little when seen for the first time, but we can discuss intuitively what it means. The goal is to avoid a situation like this:




In the figure we have an open set $V$ and a curve $\gamma$ in $V$ that surrounds a hole in $V$. The function $g(z) = 1/(z - z_0)$ is holomorphic on $V$, and

$$\oint_\gamma g(z)\,dz = \oint_\gamma \frac{1}{z - z_0}\,dz = 2\pi i \operatorname{Ind}_\gamma(z_0) = 2\pi i.$$

This means that Cauchy's Theorem does not hold for such a $V$; it has holomorphic functions with nonzero integrals over closed curves. In the picture, we have $\operatorname{Ind}_\gamma(z_0) = 1$, even though $z_0 \notin V$. What we will do is to require that $\operatorname{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$ to forbid $V$ from having holes. A slightly more general condition is to require the curve $\gamma$ to be homotopic to the zero curve, but we will not go there.

**3.9.1. Theorem. (Cauchy's Theorem and Integral Formula.)**

*Let $V \subset \mathbb{C}$ be open, $\gamma$ a closed curve in $V$ such that $\operatorname{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$, and $f$ holomorphic on $V$. Then*

$$\oint_\gamma f(z)\,dz = 0,$$

*and, for any $z \in V \setminus \gamma$,*

$$f(z)\operatorname{Ind}_\gamma(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}\,dw.$$

*The case of greatest interest occurs when $\gamma$ is simple enough—most often a circle—so that $\operatorname{Ind}_\gamma(z) = 1$. In that case,*

$$f(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}\,dw. \tag{3.14}$$

*Proof.* Define

$$g(z,w) = \begin{cases} \frac{f(w)-f(z)}{w-z}, & w \in V \setminus \{z\} \\ f'(z), & w = z \end{cases}$$

Then $g$ is continuous on $V \times V$; indeed, at $(z,w)$ with $z \neq w$, this follows directly from the continuity of $f$. And when $z = w$, if $w_n \to z$ and $z_n \to z$, then

$$\frac{f(w_n) - f(z_n)}{w_n - z_n} - f'(z) = f'(z_n) + o(w_n - z_n) - f'(z),$$

and the limit is zero by the continuity of $f'$ (by Corollary 3.5.2, since it is holomorphic). A similar estimate shows that the function $k(z) = g(z,w)$ is holomorphic for all $z$ and $w$; indeed, when $z \neq w$ we have that $k$ is a quotient of a holomorphic function by a nonzero holomorphic function, so holomorphic. And when $z = w$, we have

$$\begin{aligned}
\frac{k(w+h)-k(w)}{h} &= \frac{1}{h}\Big[\frac{f(w)-f(w+h)}{-h} - f'(w)\Big] = \frac{1}{h}\Big[\frac{f(w+h)-f(w)}{h} - f'(w)\Big] \\
&= \frac{1}{h^2}\big[f(w) + hf'(w) + \frac{h^2}{2}f''(w) + o(h^3) - f(w) - hf'(w)\big] \\
&= \frac{f''(w)}{2} + o(h) \xrightarrow[h\to 0]{} \frac{f''(w)}{2},
\end{aligned}$$

showing that $k$ is holomorphic for all $z \in V$. Define

$$F(z) = \frac{1}{2\pi i}\oint_\gamma g(z,w)\,dw, \qquad z \in V.$$

The continuity of $g$ guarantees that $F$ is defined. If $z_n \to z$, as $\gamma$ is compact we have that $g$ is continuous on the compact set $(\{z_n\} \cup \{z\}) \times \gamma$, so uniformly continuous. This guarantees that $F(z_n) \to F(z)$, and so $F$ is continuous. For a fixed $z_0 \in V \setminus \gamma$, since $\gamma$ is compact there exists a disk $z_0 + r\mathbb{D} \subset V \setminus \gamma$. Let $\Delta$ be a closed triangle inside $z_0 + r\mathbb{D}$. We have

$$\oint_\Delta F(z)\,dz = \frac{1}{2\pi i}\oint_\gamma \oint_\Delta g(z,w)\,dz\,dw = 0,$$

where the exchange of integrals is valid due to $g$ being bounded (continuous on the compact set $\Delta \times \gamma$), and the inner integral is equal to zero by Theorem 3.5.1 since $z_0 + r\mathbb{D}$ is open and convex, and as shown above $k : z \longmapsto g(z,w)$ is holomorphic. Now Morera's Theorem (3.5.3) shows that $F$ is holomorphic on $z_0 + r\mathbb{D}$. As $z_0$ was any point not in $\gamma$, we get that $F$ is holomorphic on $V \setminus \gamma$.

Let $V_0 = \{z \in \mathbb{C} \setminus \gamma : \ \mathrm{Ind}_\gamma(z) = 0\}$ and define

$$G(z) = \begin{cases} \dfrac{1}{2\pi i}\displaystyle\oint_\gamma \frac{f(w)}{w-z}\,dw, & z \in V_0 \\ \\ F(z), & z \in V \end{cases}$$

This is well defined because, if $z \in V_0 \cap V$, then

$$\frac{1}{2\pi i}\oint_\gamma \frac{f(z)}{z-w}\,dw = f(z)\,\mathrm{Ind}_\gamma(z) = 0$$

and so

$$F(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)-f(z)}{z-w}\,dw = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{z-w}\,dw.$$

As our hypothesis is that $\mathrm{Ind}_\gamma(z) = 0$ on $\mathbb{C} \setminus V$, we have $\mathbb{C} \setminus V \subset V_0$, and so $V \cup V_0 = \mathbb{C}$ and $G$ is entire. When $z$ is outside of $(r+1)\mathbb{D}$ and $\gamma \subset r\mathbb{D}$,

$$|G(z)| = \left|\frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}\,dw\right| \le \frac{\max\{|f(w)| : \ w \in \gamma\}\,k}{|z|-r} \le \max\{|f(w)| : \ w \in \gamma\}\,k,$$

where $k > 0$ satisfies $|\gamma'(t)| \le k$ for all $t$ (note that $\gamma$ is by hypothesis piecewise differentiable, so $\gamma'$ is continuous on a finite number of intervals and hence bounded). It follows that $G$ is an entire bounded function. By Liouville's Theorem (Corollary 3.7.4) we get that $G(z) = 0$ and so $F(z) = 0$ for all $z \in V$. Then for $z \in V \setminus \gamma$

$$0 = F(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)-f(z)}{w-z}\,dw = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}\,dw - f(z)\,\mathrm{Ind}_\gamma(z),$$

giving us Cauchy's Integral

$$f(z)\,\mathrm{Ind}_\gamma(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}\,dw.$$

If we now apply the formula to the holomorphic function $g(z) = (z - z_0)f(z)$, where $z_0 \in V \setminus \gamma$,

$$0 = g(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{g(z)}{z - z_0}\, dz = \frac{1}{2\pi i} \oint_\gamma \frac{(z - z_0)f(z)}{z - z_0}\, dz = \frac{1}{2\pi i} \oint_\gamma f(z)\, dz. \qquad \square$$

**3.9.2. Corollary.**

*Let $V \subset \mathbb{C}$ be open, $\gamma_1$ and $\gamma_2$ closed curves, and $f$ holomorphic on $V$. If $\operatorname{Ind}_{\gamma_1}(z) = \operatorname{Ind}_{\gamma_2}(z)$ for all $z \in \mathbb{C} \setminus V$, then*

$$\oint_{\gamma_1} f(z)\, dz = \oint_{\gamma_2} f(z)\, dz.$$

*Proof.* Apply Theorem 3.9.1 to the curve $\gamma_1 - \gamma_2$, noting that $\operatorname{Ind}_{\gamma_1 - \gamma_2}(z) = \operatorname{Ind}_{\gamma_1}(z) - \operatorname{Ind}_{\gamma_2}(z)$. $\square$

**3.9.3. Corollary.**

*Let $V \subset \mathbb{C}$ be open and $f$ holomorphic on $V$ with zeros $z_1, \ldots, z_m$ (counting multiplicities). If $\gamma$ is a closed curve on $V$ with $\operatorname{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$ and $\{z_1, \ldots, z_n\} \cap \gamma = \varnothing$, then*

$$\oint_\gamma \frac{f'(z)}{f(z)}\, dz = 2\pi i \sum_{k=1}^{m} \operatorname{Ind}_\gamma(z_k).$$

*Proof.* We can repeat the proof of Corollary 3.6.5 but now using Theorem 3.9.1 instead of Theorem 3.5.1. $\square$

Below we state Corollary 3.9.4 just for curves with index 1, as that is the common way it is used. The reader should have no difficulty stating and proving the general version.

**3.9.4. Corollary. *(Generalized Cauchy Integral Formula)***

*Let $V \subset \mathbb{C}$ be open and convex, $z \in V$, $\gamma$ a closed curve in $V$ with $\operatorname{Ind}_\gamma(z) = 1$, and $f$ holomorphic on $V$. Then $f$ is infinitely differentiable at $z$ and*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_\gamma \frac{f(w)}{(w - z)^{n+1}}\, dw. \tag{3.15}$$

*Proof.* We use induction. By Theorem 3.9.1 we can express $f(z)$ using (3.10); this is the case $n = 0$ in (3.15). So we assume that (3.15) holds for $n$. Then

$$\frac{f^{(n)}(z+h) - f^{(n)}(z)}{h} = \frac{n!}{2\pi i} \oint_\gamma f(w) \left[ \frac{1}{h(w-z-h)^{n+1}} - \frac{1}{h(w-z)^{n+1}} \right] dw$$
$$= \frac{n!}{2\pi i} \oint_\gamma f(w) \frac{\sum_{k=0}^{n} (w-z-h)^k (w-z)^{n-k}}{(w-z-h)^{n+1}(w-z)^{n+1}} \, dw.$$

If $h$ is small enough so that $z + h$ is far from $\gamma$, the integrand in the last integrand is bounded, and Dominated Convergence applies. Then

$$f^{(n+1)}(z) = \frac{n!}{2\pi i} \oint_\gamma f(w) \frac{(n+1)\,(w-z)^n}{(w-z)^{2n+2}} \, dw = \frac{(n+1)!}{2\pi i} \oint_\gamma \frac{f(w)}{(w-z)^{n+2}} \, dw.$$

This completes the inductive step, and (3.15) is proven. □

Let us put (3.15) to use, to establish an estimate that will be of use later. The subtlety here is that we want a bound in terms of the real part of $f$. Had we wanted an estimate with the max of $f$, the argument is a one-liner by using (3.15).

***3.9.5. Proposition.***

*Let $R, \varepsilon > 0$ and $f$ be holomorphic on $(R+\varepsilon)\mathbb{D}$. Then*

$$|f^{(n)}(0)| \le \frac{2n!}{R^n} \max \left\{ \operatorname{Re}\big(f(z) - f(0)\big) : \ |z| = R \right\}. \tag{3.16}$$

*Proof.* Assume initially that $f(0) = 0$. Write $f^{(n)}(0) = e^{i\theta} |f^{(n)}(0)|$. On the curve $\gamma = R\mathbb{T}$, since $f(z) z^{n-1}$ is holomorphic,

$$0 = \frac{1}{R^n} \int_\gamma f(z) z^{n-1} \, dz = \int_0^{2\pi} f(Re^{it}) \, e^{int} \, dt. \tag{3.17}$$

Similarly,

$$0 = f(0) = \int_\gamma f(z) z^{-1} \, dz = \int_0^{2\pi} f(Re^{it}) \, dt. \tag{3.18}$$

Using (3.15),

$$R^n \frac{|f^{(n)}(0)|}{n!} = \frac{R^n}{2\pi i} \int_\gamma \frac{e^{-i\theta} f(z)}{z^{n+1}} \, dz = \int_0^{2\pi} f(Re^{it}) \, e^{-int+\theta} \, dt. \tag{3.19}$$

Multiplying (3.17) by $e^{i\theta}$ and adding one half of (3.17) and (3.19) to (3.18),

$$R^n \frac{|f^{(n)}(0)|}{2n!} = \int_0^{2\pi} f(Re^{it}) \, (1 + \cos(nt + \theta)) \, dt.$$

As the left-hand-side is real, we can replace the right-hand-side with its real part. Then

$$R^n \frac{|f^{(n)}(0)|}{2n!} = \int_0^{2\pi} \text{Re } f(Re^{it})\,(1+\cos(nt+\theta)\,dt \le \max\{Ref(z):\ |z|=R\},$$

since $1+\cos(nt+\theta) \ge 0$ and its integral is 1. For arbitrary $f$, we can do the above for $f - f(0)$ and (3.16) follows. □

***3.9.6. Corollary.***

*Let $R, \varepsilon > 0$ and $f$ be holomorphic on $(R+\varepsilon)\mathbb{D}$ and $r$ with $0 < r < R$. Then*

$$\begin{aligned} \max\{|f(z)|:\ |z|=r\} \le \frac{2r}{R-r} \max\left\{\operatorname{Re} f(z):\ |z|=R\right\} \\ + \frac{R+r}{R-r}\,|f(0)|. \end{aligned} \tag{3.20}$$

*Proof.* Assume first that $f(0)=0$. Write $M_{f,R} = \max\{\operatorname{Re} f(z):\ |z|=R\}$. When $|z|=r$ we have, using Proposition 3.9.5,

$$|f(z)| \le \sum_{n=1}^{\infty} \frac{|f^{(n)}(0)|}{n!}\,|z|^n \le 2\,M_{f,R} \sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n = \frac{2r}{R-r}\,M_{f,R}.$$

For arbitrary $f$,

$$\begin{aligned} |f(z)| &\le |f(z)-f(0)| + |f(0)| \le \frac{2r}{R-r}\,M_{f-f(0),R} + |f(0)| \\ &\le \frac{2r}{R-r}\,M_{f,R} + \frac{2r}{R-r}\,|f(0)| + |f(0)| = \frac{2r}{R-r}\,M_{f,R} + \frac{R+r}{R-r}\,|f(0)|. \end{aligned}$$

□

**3.9.1. Exercises.**

(3.9.1) Let $f$ be entire and such that there exist $c, r > 0$ and $n \in \mathbb{N}$ such that $|f(z)| \le c|z|^n$ for all $z$ with $|z| > r$. Prove that $f$ is a polynomial with $\deg f \le n$.

(3.9.2) Let $f$ be entire and such that, for $z$ big enough, $\operatorname{Re} f(z) \le c|z|^s$ for some $c > 0$ and $s > 0$. Show that $f$ is a polynomial of degree at most $m = \lfloor s \rfloor$.

## 3.10. Meromorphic Functions and Residues

We have repeatedly seen the nice behaviour of holomorphic functions. We will now see that in many cases, studying the points where a function fails to be holomorphic can provide great insight.

**3.10.1. Definition.**

*If $V \subset \mathbb{C}$ is open, $z_0 \in V$, and $f$ is holomorphic on $V \setminus \{z_0\}$, we say that $f$ has a **singularity** at $z_0$. The singularity is said to be **removable** if $f$ can be extended to $z_0$ in such a way that it is holomorphic on $V$. The singularity is a **pole of order** $n$ (where $n \in \mathbb{N}$) if $(z-z_0)^n f(z)$ has a removable singularity at $z_0$ and it takes a nonzero value at $z_0$. A singularity that is neither removable nor a pole is **essential**.*

The above are the only three possible behaviours for a singularity:

***3.10.2. Proposition.***

*Let $V \subset \mathbb{C}$ be open, $z_0 \in V$, $f$ holomorphic on $V \setminus \{z_0\}$. Then precisely one of the following three conditions occur:*

*(i) $f$ has a removable singularity at $z_0$;*

*(ii) $f$ has a pole at $z_0$;*

*(iii) for any $r > 0$ such that $z_0 + r\,\mathbb{D} \subset V$, the image $f\big((z_0 + r\,\mathbb{D}) \setminus \{z_0\}\big)$ is dense in $\mathbb{C}$.*

*When $f$ is bounded on some disk centered at $z_0$, the singularity is removable.*

*Proof.* Consider first the case where $f$ is bounded on some disk $z_0 + r\mathbb{D} \subset V$. So $|f(z)| \leq c$ whenever $|z - z_0| < r$. Define

$$g(z) = \begin{cases} 0, & z = z_0 \\ (z-z_0)^2 f(z), & z \in V \setminus \{z_0\} \end{cases}$$

For every point other than $z_0$, the function $g$ is holomorphic since it is a product of holomorphic functions. At $z_0$, we have

$$\left|\frac{g(z) - g(z_0)}{z - z_0}\right| = |z - z_0|\,|f(z)| \leq c\,|z - z_0| \to 0,$$

so $g'(z_0) = 0$. Then $g$ is holomorphic on $V$, thus analytic by Corollary 3.5.2. As the disk $z_0 + r\,\mathbb{D}$ is inside $V$, we can write

$$g(z) = \sum_{k=2}^{\infty} a_k (z - z_0)^k, \qquad z \in z_0 + r\,\mathbb{D}.$$

Note that $a_0 = g(z_0) = 0$, $a_1 = g'(z_0) = 0$. Defining $f(z_0) = a_2$, we obtain that

$$f(z) = \sum_{k=0}^{\infty} a_{k+2}\,(z - z_0)^k, \qquad z \in z_0 + r\,\mathbb{D}$$

and this is the holomorphic extension of $f$. So the singularity is removable.

Now the case where $f$ is unbounded on any disk around $z_0$. If (iii) occurs, there is nothing to be done. So we assume that it fails. That is, there exists $r > 0$, $\varepsilon > 0$, and $w \in \mathbb{C}$ such that $|f(z) - w| \geq \varepsilon$ for all $z \in (z_0 + r\,\mathbb{D}) \setminus \{z_0\}$. Let

$$g(z) = \frac{1}{f(z) - w}, \qquad z \in (z_0 + r\,\mathbb{D}) \setminus \{z_0\}.$$

The conditions above guarantee that the denominator stays away from 0, and so $g$ is holomorphic on $(z_0 + r\,\mathbb{D}) \setminus \{z_0\}$. Also $|g(z)| \leq \frac{1}{\varepsilon}$ for all $z$ in the punctured disk. By the first part of the proof, the singularity of $g$ at $z_0$ is removable; that is, $g$ extends to a holomorphic function on $z_0 + r\,\mathbb{D}$. If $g(z_0) \neq 0$, then by continuity $|f(z) - w| \leq \frac{2}{|g(z_0)|}$ on some disk $z_0 + s\,\mathbb{D}$; so $f$ is bounded in such disk, and the singularity is removable by the first part of the proof.

When $g(z_0) = 0$, by Corollary 3.6.3 we can write $g(z) = (z - z_0)^n h(z)$ with $n \geq 1$, $h$ analytic on $z_0 + r\,\mathbb{D}$ and $h(z_0) \neq 0$. Since $g(z) \neq 0$ for all $z \in (z_0 + r\,\mathbb{D}) \setminus \{z_0\}$ we get that $h(z) \neq 0$ in $z_0 + r\,\mathbb{D}$. So $1/h$ is analytic in $z_0 + r\,\mathbb{D}$. Thus

$$f(z) - w = \frac{1}{g(z)} = \frac{1}{(z - z_0)^n}\,\frac{1}{h(z)}$$

and so $(z - z_0)^n f(z) = (z - z_0)^n w + \frac{1}{h(z)}$ has a removable singularity at $z_0$. We have obtained (ii). $\square$

When (ii) in Proposition 3.10.2 occurs, by writing $(z - z_0)^n f(z)$ as a power series we see that the condition is equivalent to the existence of (unique!) numbers $a_1, \ldots, a_n \in \mathbb{C}$, with $a_n \neq 0$, such that

$$f(z) - \sum_{k=1}^{n} \frac{a_k}{(z - z_0)^k} \tag{3.21}$$

has a removable singularity at $z_0$ (Exercise 3.10.1). The expression after the minus sign in (3.21) is the **principal part** of $f$. This is also equivalent to saying that $f$ has a **Laurent expansion**

$$f(z) = \sum_{k=-n}^{\infty} c_k (z - z_0)^k. \tag{3.22}$$

**3.10.3. Definition.**

*Let $V \subset \mathbb{C}$ be open. A function $f : V \to \mathbb{C}$ is* ***meromorphic*** *if there exists $P_0 \subset V$, with no accumulation points, such that $f$ is holomorphic on $V \setminus P_0$, and each $z \in P_0$ is a pole for $f$. For each $z_0 \in P_0$, if $n$ is the order of $z_0$ then the* ***residue*** *of $f$ at $z_0$ is the number*

$$\operatorname{Res}(f, z_0) = \frac{1}{(n-1)!} \lim_{z \to z_0} \frac{d^{n-1}}{dz^{n-1}} (z - z_0)^n f(z). \tag{3.23}$$

The requirement that the poles do not accumulate is an artificial but sensible restriction that removes badly behaving functions.

When the meromorphic function is expressed around its pole $z_0$ by its Laurent series as in (3.22), its residue at $z_0$ is the coefficient $c_{-1}$. In (3.21), we have $\operatorname{Res}(f, z_0) = a_1$.

When $f$ is given explicitly, it is often not hard to determine the residues and there is no need to calculate iterated derivatives. For instance, if $f(z) = \frac{1}{(z+1)^2(z-3)}$, then using partial fractions we can write

$$f(z) = -\frac{1}{16(z+1)} - \frac{1}{4(z+1)^2} + \frac{1}{16(z-3)}.$$

Now we can immediately see that

$$\operatorname{Res}(f, -1) = -\frac{1}{16}, \qquad \operatorname{Res}(f, 3) = \frac{1}{16}.$$

If instead we had $f(z) = \frac{e^z}{(z+1)^2(z-3)}$, now we cannot read the series' coefficients out of the partial fractions formula as above. But it is fairly easy to use (3.23): since $-1$ is a pole of order 2,

$$\operatorname{Res}(f, -1) = \frac{d}{dz} \frac{(z+1)^2 e^z}{(z+1)^2(z+3)}\bigg|_{z=-1} = \frac{d}{dz} \frac{e^z}{z-3}\bigg|_{z=-1} = \frac{e^z(z-4)}{(z-3)^2}\bigg|_{z=-1} = -\frac{5}{16e}.$$

We care about residues because of this generalization of Cauchy's Integral Formula:

***3.10.4. Theorem. (Residue Theorem)***

*Let $V \subset \mathbb{C}$ be open, and $f$ meromorphic on $V$ with poles $P_0$. Let $\gamma$ be a closed curve in $V$ such that $P_0 \cap \gamma = \varnothing$ and $\operatorname{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$.*

*Then*

$$\oint_\gamma f(w)\,dw = 2\pi i \sum_{z\in P_0} \mathrm{Res}(f,z)\ \mathrm{Ind}_\gamma(z). \tag{3.24}$$

*Proof.* Recall that, by definition, $P_0$ has no accumulation points. Let $P_1 = \{z \in P_0 : \mathrm{Ind}_\gamma(z) \neq 0\}$. By Proposition 3.4.2 we have that $\mathrm{Ind}_\gamma(z)$ is constant on each connected component of $\mathbb{C}\setminus\gamma$. The condition $\mathrm{Ind}_\gamma(z) = 0$ on $\mathbb{C}\setminus V$ guarantees that $\mathbb{C}\setminus V$ is contained in the unique unbounded connected component of $\mathbb{C}\setminus\gamma$. Thus $P_1$ is bounded; having no accumulation points, it is finite. So the sum in (3.24) is always finite; it may be zero if $P_1 = \varnothing$, in which case we recover Theorem 3.9.1.

Using Proposition 3.10.2 and (3.21) we may write, for each $z_1,\dots,z_m \in P_1$

$$f(z) = g_r(z) + \sum_{k=1}^{n_r} \frac{a_{r,k}}{(z-z_r)^k}, \qquad r = 1,\dots,m,$$

where $g_r$ is holomorphic at $z_r$ and $\mathrm{Res}(f,z_r) = a_{r,1}$. Then the function

$$f(z) - \sum_{r=1}^{m}\sum_{k=1}^{n_r} \frac{a_{r,k}}{(z-z_r)^k}$$

is holomorphic on $V\setminus(P_0\setminus P_1)$, with removable singularities. By Theorem 3.9.1,

$$\begin{aligned}
0 &= \oint_\gamma \Big[f(z) - \sum_{r=1}^{m}\sum_{k=1}^{n_r} \frac{a_{r,k}}{(z-z_r)^k}\Big]\,dz \\
&= \oint_\gamma f(z)\,dz - \sum_{r=1}^{m}\sum_{k=1}^{n_r} a_{r,k} \oint_\gamma \frac{1}{(z-z_r)^k}\,dz \\
&= \oint_\gamma f(z)\,dz - \sum_{r=1}^{m} a_{r,1} \oint_\gamma \frac{1}{z-z_r}\,dz \\
&= \oint_\gamma f(z)\,dz - 2\pi i \sum_{r=1}^{m} \mathrm{Res}(f,z_r)\ \mathrm{Ind}_\gamma(z_r).
\end{aligned}$$

There are many ways to see that the integrals $\oint_\gamma \frac{1}{(z-z_r)^k}\,dz$ are zero for $k>1$. Among them is direct calculation, and also Corollary 3.9.4 applied to the constant function 1. ∎

Let us now check a few examples of how Theorem 3.10.4 can be used.

*3.10.5. Example.* Let $\gamma$ be the circle of radius 4, centered at the origin. Let us calculate
$$\int_\gamma \frac{1}{(z^2+9)(z+9)}\,dz.$$
One cannot apply Cauchy's formula directly because the function is not holomorphic inside the disk (since $z^2+9=(z+3i)(z-3i)$). Using partial fractions we can write
$$f(z)=\frac{1}{(z^2+9)(z+9)}=\frac{\alpha}{z+3i}+\frac{\beta}{z-3i}+\frac{\gamma}{z+9},$$
where
$$\alpha=-\frac{1}{180}+\frac{i}{60},\quad \beta=-\frac{1}{180}-\frac{i}{60},\quad \gamma=\frac{1}{90}.$$
Since only $\pm 3i$ are inside the curve, we only need those two residues. Both points are poles of order 1 and their index is 1, so all we need is the coefficients of $(z\pm 3i)^{-1}$ in the expansion. It is apparent then that
$$\operatorname{Res}(f,-3i)=-\frac{1}{180}+\frac{i}{60},\qquad \operatorname{Res}(f,3i)=-\frac{1}{180}-\frac{i}{60}.$$
Then, by Theorem 3.9.1,
$$\int_\gamma \frac{1}{(z^2+9)(z+9)}\,dz=2\pi i\Big(-\frac{1}{180}-\frac{i}{60}-\frac{1}{180}+\frac{i}{60}\Big)=-\frac{\pi i}{45}.$$
Let us clarify that this integral, after having decomposed $f$ using partial fractions, can be done also by Cauchy's Theorem. So we will move to more substantial examples of using residues.

*3.10.6. Example. The Series $\sum_{n>0}\frac{1}{n^2}$.* Consider the function $f(z)=\frac{\cos\pi z}{z^2\,\sin\pi z}$. This function is meromorphic, with poles at all integers $z$. These poles are all of order one: to see this, when $n\neq 0$ we have
$$\begin{aligned}\frac{(z-n)\cos\pi z}{z^2\,\sin\pi z}&=\frac{i(z-n)(e^{i\pi z}+e^{-i\pi z})}{z^2\,(e^{i\pi z}-e^{-i\pi z})}=\frac{i(z-n)(1+e^{-2i\pi z})}{z^2\,(1-e^{-2i\pi z})}\\&=\frac{i(z-n)(1+e^{-2i\pi(z-n)})}{z^2\,(1-e^{-2i\pi(z-n)})}=\frac{i(z-n)(1+e^{-2i\pi(z-n)})}{z^2\,(2i\pi(z-n)+o(z-n)^2)}\\&=\frac{1+e^{-2i\pi(z-n)}}{(2z^2\pi+o(z-n))}\xrightarrow[z\to n]{}\frac{1}{n^2\pi}\end{aligned}$$
(if the poles had been of higher order, the limits would have not existed). Then
$$\operatorname{Res}(f,n)=\frac{1}{n^2\pi}.$$
When $n=0$ we have
$$\frac{z^3\,\cos\pi z}{z^2\,\sin\pi z}\xrightarrow[z\to 0]{}\frac{1}{\pi}.$$

This shows that the pole at 0 has order 3. Instead of calculating a second derivative, and using a geometric expansion on the third equality,

$$\begin{aligned} f(z) &= \frac{1-\frac{\pi^2 z^2}{2}+o(z^4)}{z^2\left(\pi z-\frac{\pi^3 z^3}{6}+o(z^5)\right)} = \frac{1}{\pi z^3}\,\frac{1-\frac{\pi^2 z^2}{2}+o(z^4)}{1-\frac{\pi^2 z^2}{6}+o(z^4)} \\ &= \frac{1}{\pi z^3}\left(1-\frac{\pi^2 z^2}{2}+o(z^4)\right)\left(1+\frac{\pi^2 z^2}{6}+o(z^4)\right) \\ &= \frac{1}{\pi z^3}-\frac{\pi}{3z}+o(z). \end{aligned}$$

Thus

$$\operatorname{Res}(f,0) = -\frac{\pi}{3}.$$

We see that for $n \neq 0$ it was easier to calculate the residue by definition, as otherwise we would have to calculate the series expansion in terms of $z-n$, which is easy to do for the exponentials as above, but we would have to deal with the $z^2$ in the denominator. At $z=0$ this was no issue, so obtaining the expansion was easier.

For $N \in \mathbb{N}$, let $\gamma_N(t)$ describe the square with vertices $(N+\frac{1}{4})(\pm 1 \pm i)$. By Theorem 3.10.4,

$$\frac{1}{2\pi i}\oint_{\gamma_N}\frac{\cos \pi z}{z^2\,\sin \pi z}\,dz = -\frac{\pi}{3}+\sum_{0<|n|\le N}\frac{1}{n^2\pi} = -\frac{\pi}{3}+2\sum_{0<n<N}\frac{1}{n^2\pi}. \tag{3.25}$$

The curve $\gamma_N$ is made of four segments. We can estimate

$$\left|\frac{\cos \pi z}{\sin \pi z}\right| = \left|1+\frac{2}{e^{-2i\pi z}-1}\right| \le 1+\frac{2}{|e^{-2i\pi z}-1|}.$$

On the two vertical segments, we have $z=(N+\frac{1}{4})(\pm 1+it)$, $t\in[-1,1]$. Then

$$\begin{aligned} |e^{-2i\pi z}-1| &= |e^{2\pi t(N+\frac{1}{4})}e^{\mp 2i\pi(N+\frac{1}{4})}-1| = |e^{2\pi t(N+\frac{1}{4})}-e^{\pm 2i\pi(N+\frac{1}{4})}| \\ &\ge |\sin(2\pi N+\tfrac{\pi}{2})| = 1, \end{aligned}$$

the inequality coming from $|w| \ge |\operatorname{Im} w|$. On the horizontal segment $z=(N+\frac{1}{4})\,(t-i)$,

$$\begin{aligned} |e^{-2i\pi z}-1| &= |e^{-2\pi i t(N+\frac{1}{4})}e^{-2\pi(N+\frac{1}{4})}-1| = |e^{2i\pi t(N+\frac{1}{4})}-e^{-2\pi N}e^{-\pi/2}| \\ &\ge 1-e^{-2\pi N}e^{-\pi/2} \ge 1-e^{-\frac{\pi}{2}}. \end{aligned}$$

On the horizontal segment $z=(N+\frac{1}{4})\,(t+i)$,

$$\begin{aligned} |e^{-2i\pi z}-1| &= |e^{-2\pi i t(N+\frac{1}{4})}e^{2\pi(N+\frac{1}{4})}-1| = |e^{2\pi(N+\frac{1}{4})}-e^{-2\pi i t(N+\frac{1}{4})}| \\ &\ge e^{2\pi(N+\frac{1}{4})}-1 \ge e^{\pi/2}-1. \end{aligned}$$

As $e^{\frac{\pi}{2}} > 4$, we have that $e^{\pi/2} - 1 > 1 - e^{-\pi/2}$. We conclude that on all four segments of $\gamma_N$ we have $|e^{-2i\pi z} - 1| \geq 1 - e^{-\pi/2}$. Therefore, on $\gamma_N$,

$$\left|\frac{\cos \pi z}{\sin \pi z}\right| \leq 1 + \frac{2}{1 - e^{-\frac{\pi}{2}}} = \frac{3 - e^{-\frac{\pi}{2}}}{1 - e^{-\frac{\pi}{2}}}.$$

Since the length of $\gamma_N$ is $8N + 2$, and using Lemma 3.3.4,

$$\oint_{\gamma_N} \left|\frac{\cos \pi z}{z^2 \sin \pi z}\right| dz \leq \frac{(3 - e^{-\frac{\pi}{2}})(8N + 2)}{(1 - e^{-\frac{\pi}{2}})N^2}, \tag{3.26}$$

and so the integral in (3.25) goes to zero as $N \to \infty$. Then in the limit (3.25) becomes

$$0 = -\frac{\pi}{3} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n^2},$$

giving the known result $\sum_n \frac{1}{n^2} = \frac{\pi^2}{6}$.

The same idea can be pushed for any even power. If we consider instead $f(z) = \frac{\cos \pi z}{z^4 \sin \pi z}$, the computation for the residues at $n \neq 0$ is the same, with an extra $n^2$ in the denominator; so all we need to calculate is the residue at zero. We have

$$\begin{aligned}
f(z) &= \frac{1 - \frac{\pi^2 z^2}{2} + \frac{\pi^4 z^4}{24} + o(z^6)}{z^4 \left(\pi z - \frac{\pi^3 z^3}{6} + \frac{\pi^5 z^5}{120} + o(z^7)\right)} = \frac{1}{\pi z^5} \frac{1 - \frac{\pi^2 z^2}{2} + \frac{\pi^4 z^4}{24} + o(z^6)}{1 - \frac{\pi^2 z^2}{6} + \frac{\pi^4 z^4}{120} + o(z^6)} \\
&= \frac{1}{\pi z^5} \left(1 - \frac{\pi^2 z^2}{2} + \frac{\pi^4 z^4}{24} + o(z^6)\right)\left(1 + \frac{\pi^2 z^2}{6} - \frac{\pi^4 z^4}{120} + \left(\frac{\pi^2 z^2}{6}\right)^2 + o(z^6)\right) \\
&= \frac{1}{\pi z^5} - \frac{\pi}{3z^3} - \frac{\pi^3}{45z} + o(z).
\end{aligned}$$

Thus

$$\mathrm{Res}(f, 0) = -\frac{\pi^3}{45}.$$

Then, as above, we get

$$0 = -\frac{\pi^3}{45} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n^4},$$

from where we get

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

More generally, with the same idea, one can show that

$$\mathrm{Res}\left(\frac{\cos \pi z}{z^{2k} \sin \pi z}, 0\right) = \frac{(2\pi i)^{2k} B_{2k}}{\pi\,(2k)!}, \tag{3.27}$$

where $B_{2k}$ are the **Bernoulli numbers**. It follows that

$$\sum_{k=1}^{\infty} \frac{1}{n^{2k}} = \frac{(-1)^{k+1}(2\pi)^{2k} B_{2k}}{2(2k)!}$$

The Bernoulli numbers are given by

$$B_n = \sum_{k=0}^{n} \sum_{h=0}^{k} (-1)^h \binom{k}{h} \frac{h^n}{k+1}.$$

They appear as the coefficients in the expansion

$$\frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k\, z^k}{k!};$$

that is, with $g(z) = \frac{z}{e^z-1}$,

$$B_k = g^{(k)}(0).$$

The function $\frac{\cos \pi z}{\sin \pi z}$ is useful here because it is meromorphic with poles on the integers and residue $\frac{1}{\pi}$ at each of them. A fact that is not useful to find the residues as above but interesting nonetheless, is the equality

$$\frac{\cos \pi z}{\sin \pi z} = \frac{1}{\pi z} + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{2z}{z^2 - k^2} \tag{3.28}$$

(see Example 3.10.10). Surprisingly, there is no obvious extension of the above techniques to the case where $k$ is odd. The fact that $\sum_k \frac{1}{k^3}$ is irrational was only proven by Apéry in 1978 [**Apé79**, **VdPA79**], more than two hundred years after Euler's early results. It is still unknown whether it is transcendental.

The ideas above work for a great number of functions beyond $f(z) = z^{-2n}$. We refer to Exercises 3.10.13 to 3.10.15 and 3.10.17 to 3.10.19.

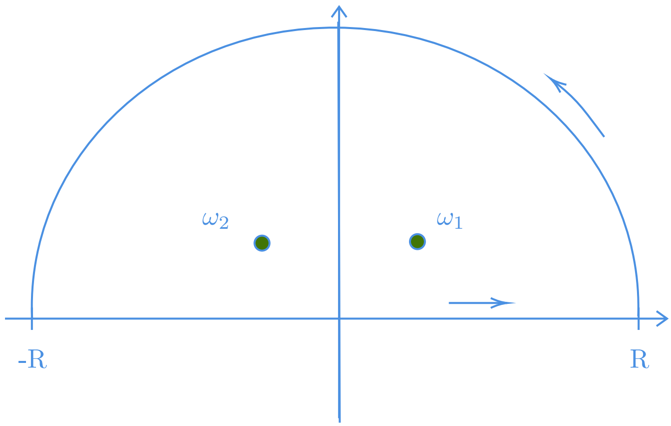
*3.10.7. Example.* Let us calculate

$$\int_{-\infty}^{\infty} \frac{x^2}{1+x^4}\, dx.$$

The meromorphic function $f(z) = \frac{z^2}{1+z^4}$ has four poles at

$$\omega_1 = \frac{1+i}{\sqrt{2}}, \quad \omega_2 = \frac{-1+i}{\sqrt{2}}, \quad \omega_3 = \frac{-1-i}{\sqrt{2}}, \quad \omega_4 = \frac{1-i}{\sqrt{2}},$$

the four fourth roots of $-1$.




Over the semicircle $\gamma_{2,R}(t) = R\, e^{it}$, $t \in [0, \pi]$, we have

$$\left| \int_{\gamma_2} \frac{z^2}{1+z^4}\, dz \right| = \left| \int_0^{\pi} \frac{R^2 e^{2it}\, Ri\, e^{it}}{1 + R^4 e^{4it}}\, dt \right| \le \frac{1}{R} \int_0^{\pi} \frac{1}{|1/R^4 + e^{it}|}\, dt$$

$$\le \frac{\pi}{R\left(1 - \frac{1}{R^4}\right)} \xrightarrow[R\to\infty]{} 0.$$

Thus, by Theorem 3.10.4,

$$\int_{-\infty}^{\infty} \frac{x^2}{1+x^4}\, dx = \lim_{R\to\infty} \int_{-R}^{R} \frac{x^2}{1+x^2}\, dx = \lim_{R\to\infty} \int_{\gamma_{1,R}} \frac{z^2}{1+z^4}\, dz$$

$$= \lim_{R\to\infty} \int_{\gamma_{1,R}+\gamma_{2,R}} \frac{z^2}{1+z^4}\, dz = 2\pi i[\operatorname{Res}(f, \omega_1) + \operatorname{Res}(f, \omega_2)].$$

For the residues, since the poles have order 1,

$$\operatorname{Res}(f, \omega_1) = \lim_{z\to\omega_1} \frac{(z-\omega_1)z^2}{1+z^4} = \frac{\omega_1^2}{(\omega_1-\omega_2)(\omega_1-\omega_3)(\omega_1-\omega_4)}$$

$$= \frac{i}{\sqrt{2}\,\sqrt{2}(1+i)\,\sqrt{2}\, i} = \frac{1-i}{4\sqrt{2}}.$$

Similarly,

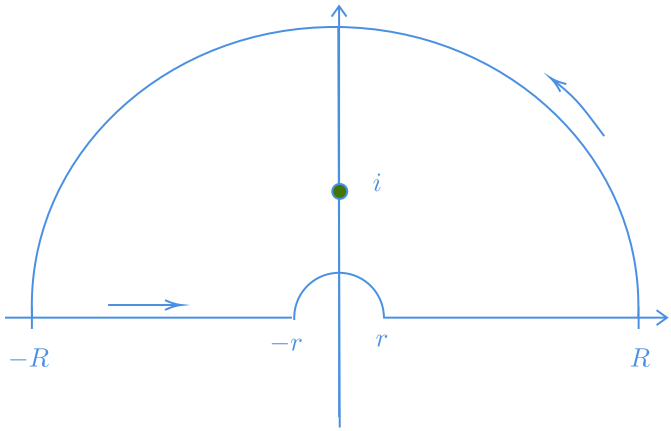
$$\operatorname{Res}(f, \omega_2) = \lim_{z\to\omega_2} \frac{(z-\omega_2)z^2}{1+z^4} = \frac{\omega_2^2}{(\omega_2-\omega_1)(\omega_2-\omega_3)(\omega_2-\omega_4)}$$

$$= \frac{-i}{-\sqrt{2}\,\sqrt{2}i\,(-1+i)\sqrt{2}} = \frac{-1-i}{4\sqrt{2}}.$$

Then

$$\int_{-\infty}^{\infty} \frac{x^2}{1+x^4}\, dx = 2\pi i\left(\frac{1-i}{4\sqrt{2}} - \frac{1+i}{4\sqrt{2}}\right) = \frac{\pi}{\sqrt{2}}.$$

*3.10.8. Example.* Let us now find
$$\int_0^\infty \frac{\log x}{(1+x^2)^2}\, dx.$$
We will use the branch of the logarithm where the argument is in $(0, 2\pi)$. The function $f(z) = \frac{\log z}{(1+z^2)^2}$ has an essential singularity at 0 and two poles at $\pm i$, both of order 2. Fix $r, R > 0$ with $r < 1 < R$. Let $\gamma$ be the curve




By Theorem 3.10.4,
$$\oint_\gamma \frac{\log z}{(1+z^2)^2}\, dz = 2\pi i \operatorname{Res}(f, i). \tag{3.29}$$
The residue is
$$\operatorname{Res}(f, i) = \lim_{z\to i} \frac{d}{dz} \frac{(z-i)^2 \log z}{(z-i)^2(z+i)^2} = \frac{\frac{1}{z}(z+i)^2 - 2(z+i)\log z}{(z+i)^4}\Big|_{z=i} = \frac{\pi}{8} + \frac{i}{4}.$$
We consider the four parts of $\gamma$ separately. In the first segment, we take $\gamma_1(t) = t$, $t \in [r, R]$. Then
$$\int_{\gamma_1} f(z)\, dz = \int_r^R \frac{\log t}{(1+t^2)^2}\, dt.$$
Second we consider $\gamma_2$ the arc with radius $R$. So $\gamma_2(t) = Re^{it}$, $t \in [0, \pi]$. Then, for $R$ big enough,
$$\Big|\int_{\gamma_2} f(z)\, dz\Big| \le \int_0^\pi \frac{|\log(Re^{it})|}{|1+R^2e^{2it}|^2} |Ri\, e^{it}|\, dt = R\int_0^\pi \frac{|\log R + it|}{|e^{-2it} + R^2|^2}\, dt$$
$$\le R \int_0^\pi \frac{|\log R + it|}{(R^2-1)^2}\, dt \le \frac{\pi R \log R + \frac{\pi^2}{2}}{(R^2-1)^2}.$$

Now $\gamma_3$ is the second straight segment, on the side with negative real part. Here $\gamma_3(t) = t - r - R$, $t \in [r, R]$. Hence, noting that $t - r - R < 0$,

$$\int_{\gamma_3} f(z)\,dz = \int_r^R \frac{\log(t-r-R)}{(1+(t-r-R)^2)^2}\,dt$$
$$= \int_r^R \frac{\log(R+r-t)+i\pi}{(1+(R+r-t)^2)^2}\,dt = \int_r^R \frac{\log t}{(1+t^2)^2}\,dt + i\pi \int_r^R \frac{1}{(1+t^2)^2}\,dt.$$

And $\gamma_4$ is the small arc with radius $r$. We can parametrize $\gamma_4(t) = re^{i(\pi-t)}$, $t \in [0, \pi]$. Then

$$\Big|\int_{\gamma_4} f(z)\,dz\Big| \le \int_0^\pi \frac{|\log(re^{i(\pi-t)})|}{|1+r^2e^{2i(\pi-r)}|^2}\,|ri\,e^{it}|\,dt = r\int_0^\pi \frac{|\log r + i(\pi-t)|}{|e^{-2i(\pi-t)}+r^2|^2}\,dt$$
$$\le r\int_0^\pi \frac{|\log r + i(\pi-t)|}{|e^{-2i(\pi-t)}+r^2|^2}\,dt \le \frac{\pi r\,\log r + \frac{\pi^2}{2}r}{(1-r^2)^2}$$

so this integral goes to zero with $r$. Now going back to (3.29) we have

$$-\frac{\pi}{2} + \frac{\pi^2 i}{4} = 2\int_r^R \frac{\log x}{(1+x^2)^2}\,dx + i\pi\int_r^R \frac{1}{(1+x^2)^2}\,dx + \int_{\gamma_2} f(z)\,dz + \int_{\gamma_4} f(z)\,dz.$$

Taking limit as $r \to 0$, $R \to \infty$, and comparing real and imaginary parts, we obtain

$$\int_0^\infty \frac{\log x}{(1+x^2)^2}\,dx = -\frac{\pi}{4}, \qquad \int_0^\infty \frac{1}{(1+x^2)^2}\,dx = \frac{\pi}{4}.$$

We finish with a remarkable characterization of meromorphic functions.

***3.10.9. Theorem. (Mittag-Leffler)***

*Let $f$ be a meromorphic function on $\mathbb{C}$ with poles $\{z_n\}$, such that $z_n \neq 0$ and $|z_n| \le |z_{n+1}|$ for all $n$. For each $n$, let $p_n(z)$ be the principal part of $f$ at $z_n$. Then there exists $g(z)$, entire, such that*

$$f(z) = g(z) + \sum_{n=1}^{\infty} \big(p_n(z) - q_n(z)\big)$$

*where $q_n$ is the Taylor expansion around $0$ of $p_n$.*

*Proof.* Fix positive integers $K_n$, to be chosen, and let

$$q_n(z) = \sum_{k=0}^{K_n} \frac{p_n^{(k)}(0)}{k!}\,z^k$$

be the expansion of order $K_n$ 0f $p_n$ at 0. Since the function $p_n(z)$ is holomorphic far from $z_n$, we can choose $K_n$ so that $|p_n(z) - q_n(z)| < 2^{-n}$ if $|z| < |z_n|/2$.

For any $z \in R\overline{\mathbb{D}} \setminus \{z_n\}$, choose $N \in \mathbb{N}$ with $|z_n| > 2R$ for all $n \geq N$. So $|z_n| > 2|z|$ for all $n \geq N$. Which means that

$$\Big|\sum_{n>N} \big(p_n(z) - q_n(z)\big)\Big| \leq \sum_{n>N} 2^{-n} = 2^{-N+1}.$$

So the series converges uniformly on any compact set that does not contain $\{z_n\}$. Now let

$$g(z) = f(z) - \sum_{n=1}^{\infty} \big(p_n(z) - q_n(z)\big).$$

The singularities of $g$ are removable since by design $p_n(z)$ will cancel the principal part at each pole. □

*3.10.10. Example.* Let $f(z) = \cot \pi z - \frac{1}{z}$. This function is meromorphic, with poles at all nonzero $n \in \mathbb{Z}$. All the poles are simple, so the principal parts are

$$p_n(z) = \frac{1}{z-n}.$$

As in the proof of Theorem 3.10.9, we choose $q_n = -1/n$, the first term of the Taylor polynomial for $p_n$. Then

$$g(z) = \frac{\pi \cos \pi z}{\sin \pi z} - \frac{1}{z} - \sum_{n \neq 0} \Big(\frac{1}{z-n} + \frac{1}{n}\Big)$$

is entire. Joining the terms with $n$ and $-n$ we get

$$g(z) = \frac{\pi \cos \pi z}{\sin \pi z} - \frac{1}{z} - \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}.$$

All three terms are odd, so $g$ is odd. In particular, $g(0) = 0$. For $N \in \mathbb{N}$, consider the circle $\gamma_N(t) = \big(N + \frac{1}{2}\big)\, e^{it}$, $0 \in [0, 2\pi)$. For $z = a + ib$, we have

$$\cos z = \cos a \cosh b + i \sin a \sinh b, \qquad \sin z = \sin a \cosh b - i \cos a \sinh b.$$

Then

$$\begin{aligned}
\Big|\frac{\cos z}{\sin z}\Big|^2 &= \frac{\cos^2 a\, \cosh^2 b + \sin^2 \sinh^2 b}{\sin^2 a \cosh^2 b + \cos^2 a \sinh^2 b} = \frac{\cos^2 a\,(1 + \sinh^2 b) + \sin^2 \sinh^2 b}{\sin^2 a(1 + \sinh^2 b) + \cos^2 a \sinh^2 b} \\
&= \frac{\cos^2 a + \sinh^2 b}{\sin^2 a + \sinh^2 b} = 1 + \frac{\cos^2 a - \sin^2 a}{\sin^2 a + \sinh^2 b} \\
&= 1 + \frac{\cos 2a}{\sin^2 a + \sinh^2 b}.
\end{aligned}$$

If $z = a + ib \in \gamma_N$ and $|a| \geq N + \frac{1}{4}$, then $N + \frac{1}{4} \leq |a| \leq N + \frac{1}{2}$, so $\sin^2 \pi a = \sin^2 |\pi a| \geq \frac{1}{2}$. Hence

$$\left|\frac{\cos \pi z}{\sin \pi z}\right|^2 = 1 + \frac{\cos 2a}{\sin^2 a + \sinh^2 b} \leq 1 + \frac{1}{\sin^2 a} \leq 3.$$

And when $|a| < N + \frac{1}{4}$, we have

$$\begin{aligned}\sinh^2 \pi b \geq \pi^2 b^2 &= \pi^2 \big(N + \frac{1}{2}\big)^2 - \pi^2 a^2 \\ &\geq \pi^2 \big(N + \frac{1}{2}\big)^2 - \pi^2 \Big(N + \frac{1}{4}\Big)^2 = \pi\Big(\frac{N}{2} + \frac{1}{4}\Big) > 1.\end{aligned}$$

Then

$$\left|\frac{\cos \pi z}{\sin \pi z}\right|^2 = 1 + \frac{\cos 2a}{\sin^2 a + \sinh^2 b} \leq 1 + \frac{1}{\sinh^2 b} \leq 2.$$

Combining the two estimates we find that for $z \in \gamma_N$,

$$\left|\frac{\cos \pi z}{\sin \pi z}\right| \leq \sqrt{3}.$$

This shows that on $\gamma_N$, using Exercise 3.10.16,

$$|g(z)| \leq \sqrt{3} + \frac{1}{N} + \sum_{n=1}^{\infty} \frac{2\Big(N + \frac{1}{2}\Big)\pi}{\Big(N + \frac{1}{2}\Big)^2 \pi^2 - n^2} = \sqrt{3} + \frac{1}{N} + \frac{1}{\Big(N + \frac{1}{2}\Big)\pi} + \pi.$$

So $g$ is entire and bounded, and hence constant by Liouville's Theorem (Corollary 3.7.4). As $g(0) = 0$, it follows that $g = 0$ and

$$\frac{\pi \cos \pi z}{\sin \pi z} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}.$$

We mentioned this equality in (3.28).

### 3.10.1. Exercises.

(3.10.1) Let $V \subset \mathbb{C}$ be open, $z_0 \in V$, and $f : V \to \mathbb{C}$ a function. Show that the following statements are equivalent:

(i) there exists $n \in \mathbb{N}$ such that $(z - z_0)^n f(z)$ has a removable singularity at $z_0$;

(ii) there exist $a_1, \ldots, a_n \in \mathbb{C}$, with $a_n \neq 0$, such that on some disk around $z_0$

$$f(z) - \sum_{k=1}^{n} \frac{a_k}{(z - z_0)^k}$$

has a removable singularity at $z_0$;

(iii) there exist coefficients $\{c_k\}_{k=-n}^{\infty} \subset \mathbb{C}$ such that, on some disk around $z_0$,

$$f(z) = \sum_{k=-n}^{\infty} c_k (z - z_0)^k.$$

(3.10.2) For each of the following functions, classify its singularities.

(i) $f(z) = \frac{\sin z}{z}$, $z \in \mathbb{C} \setminus \{0\}$;

(ii) $f(z) = \frac{e^z}{z^2}$, $z \in \mathbb{C} \setminus \{0\}$;

(iii) $f(z) = e^{1/z}$, $z \in \mathbb{C} \setminus \{0\}$.

(3.10.3) Use the Residue Theorem to evaluate $\displaystyle\oint_{|z|=2} \frac{1}{z^2(z-1)}\, dz$.

(3.10.4) Find $\displaystyle\oint_{|z|=3} \frac{e^z}{z^2+\pi^2}\, dz$.

(3.10.5) Find $\displaystyle\oint_{|z-i|=3} \frac{e^z}{z^2+\pi^2}\, dz$.

(3.10.6) Find $\displaystyle\int_0^{2\pi} \frac{1}{5+4\cos t}\, dt$.

(3.10.7) Compute $\displaystyle\int_{-\infty}^{\infty} \frac{1}{x^2+4x+5}\, dx$.

(3.10.8) Compute $\displaystyle\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}\, dx$.

(3.10.9) Show that if $s \in (0,1)$ then $\displaystyle\int_0^{\infty} \frac{x^{s-1}}{x+1}\, dx = \frac{\pi}{\sin \pi s}$.

(3.10.10) Evaluate $\displaystyle\int_{-\infty}^{\infty} \frac{x \sin x}{x^2+4x+20}\, dx$.

(3.10.11) Show that, for $0 < s < 1$, $\displaystyle\int_{-\infty}^{\infty} \frac{e^{sx}}{1+e^x}\, dx = \frac{\pi}{\sin s\pi}$.

(3.10.12) Show that if $-1 < \alpha < 3$, then $\displaystyle\int_0^{\infty} \frac{x^\alpha}{(1+x^2)^2}\, dx = \frac{\pi(1-\alpha)}{4\cos\frac{\alpha\pi}{2}}$.

(3.10.13) Using the ideas in Example 3.10.6 show that, for any meromorphic function $f$ with no integer poles and such that there exist $s, c > 0$ with $N^{1+s}\,|f(\gamma_N)| \le c$ for all $N$, where $\gamma_N$ is the curve from Example 3.10.6,

$$\sum_{n\in\mathbb{Z}} f(n) = -\pi \sum_{w\in P} \operatorname{Res}\Big(f(z)\frac{\cos \pi z}{\sin \pi z}, w\Big), \tag{3.30}$$

where $P$ is the set of poles of $f$.

(3.10.14) Show that $\displaystyle\sum_{n\in\mathbb{Z}} \frac{1}{\left(n-\frac{1}{2}\right)^2} = \pi^2$.

(3.10.15) Show that, for any $r>0$, $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2+r^2} = \frac{\pi}{2r} + \frac{\pi}{r(e^{2\pi r}-1)} - \frac{1}{2r^2}$.

(3.10.16) Let $r \in \mathbb{R}\setminus\mathbb{Z}$. Show that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2-r^2} = \frac{1}{2r^2} - \frac{\pi\cos\pi r}{2r\sin\pi r}$.

(3.10.17) Show that the function $\frac{1}{\sin\pi z}$ is bounded on the square with vertices $(N+\frac{1}{4})\,(\pm1\pm i)$, independently of $N$. Conclude that if $f$ is meromorphic, it has no integer poles, and there exist $s,c>0$ such that $N^{1+s}\,|f(\gamma_N)|\leq c$ for all $N$, where $\gamma_N$ is the curve from Example 3.10.6, then

$$\sum_{n\in\mathbb{Z}}(-1)^n f(n) = -\pi \sum_{w\in P} \operatorname{Res}\Big(\frac{f(z)}{\sin\pi z}, w\Big), \tag{3.31}$$

where $P$ is the set of poles of $f$.

(3.10.18) Show that for any $r\in\mathbb{R}\setminus\mathbb{Z}$ we have $\displaystyle\sum_{n\in\mathbb{Z}} \frac{(-1)^n}{(n+r)^2} = \frac{\pi^2\cos\pi r}{\sin^2\pi r}$.

(3.10.19) Find $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2}$.

(3.10.20) Consider the Gamma Function from Exercise 3.1.6. It is only defined for $\operatorname{Re} z>0$. But it can be continued analytically (doing **analytic continuation**) in the following way. We know that $\Gamma(z+1)=z\Gamma(z)$. This we can write as

$$\Gamma(z) = \frac{\Gamma(z+1)}{z},$$

which suggests a way to extend the function "to the left", first to the strip $-1<\operatorname{Re} z\leq 0$ (avoiding $z=0$), and subsequently to each strip $-(n+1)<\operatorname{Re} z\leq -n$ (avoiding $z=-n$). We have to avoid 0 because otherwise it appears on the denominator, and this makes the extension undefined on all non-positive integers. The formula, using that $\Gamma$ is holomorphic, shows that this extension is holomorphic. So we get a meromorphic function with poles at $-n+1$ for $n\in\mathbb{N}$.

(i) Show that these poles are simple.

(ii) Show that, for all $z\in\mathbb{C}\setminus\mathbb{Z}$,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\pi z}.$$

Exercise 3.10.9, via a substitution of the form $y=t/x$, should be useful.

(iii) Show that there exist $a, b > 0$ such that for all $z \in \mathbb{C}$

$$|\Gamma(z)| \leq ae^{b|z| \log |z|}.$$

(iv) Show that there exist $a, b > 0$ such that for all $z \in \mathbb{C}$

$$\frac{1}{|\Gamma(z)|} \leq ae^{b|z| \log |z|}.$$

## 3.11. Weierstrass' Factorization Theorem

The goal of this section is to establish a result similar to Corollary 3.6.4 but for entire functions. The obvious problem when looking at the analogy, is that one would need an infinite product. It turns out that such a thing is doable. This section is mostly based on [Con78, Chapter VII].

**3.11.1. Definition.**

*Let $\{z_n\} \subset \mathbb{C}$. The **infinite product** of the numbers $z_n$ is, if it exists, the complex number*

$$z = \lim_{k \to \infty} \prod_{n=1}^{k} z_n.$$

*When the limit exists, we write $z = \prod_{n=1}^{\infty} z_n$.*

Basic observation about the infinite product is that if any $z_n = 0$, then the product is zero. The converse is not true, though; for instance if $z_n = r$ for all $n$, with $|r| < 1$, then $z_n \neq 0$ for all $n$ but $\prod_n z_n = 0$. From the equality

$$\frac{\prod_{k=1}^{n+1} z_k}{\prod_{k=1}^{n} z_k} = z_n,$$

a necessary condition for the product to exist is that $z_n \to 1$. This allows us, when discussing convergence of products, to focus only in the tail and to assume without loss of generality that $\operatorname{Re} z_n > 0$ for all $n$; thus $\log z_n$ makes sense. By convention we will use the branch of the logarithm that is not defined on the negative half-line, that is where $-\pi < \theta < \pi$, and then $\log re^{i\theta} = \log r + i\theta$. The reason we care about the logarithm is that it transforms products into sums, so certain series will be available when discussing products. Indeed,

**3.11.2. Proposition.**

*Let $\{z_n\} \subset \mathbb{C}$ such that $\operatorname{Re} z_n > 0$ for all $n$. The following statements are equivalent:*

*(i)* $\sum_n \log z_n$ *converges;*

*(ii)* $\prod\limits_n z_n$ *converges to a nonzero number.*

*Proof.* (i) $\implies$ (ii) Let us write $s_n = \sum_{k=1}^n \log z_k$ the partial sum, and $s = \sum_k \log z_k$. Then $s_n \to s$. As the exponential is continuous,

$$\prod_{n=1}^{\infty} z_n = \lim_{n\to\infty} \prod_{k=1}^{n} z_k = \lim_{n\to\infty} \exp\left(\sum_{k=0}^{n} \log z_k\right) = \exp\left(\sum_{k=0}^{\infty} \log z_k\right) = e^s.$$

So the infinite product exists and it is equal to $\exp\left(\sum\limits_{k=0}^{\infty} \log z_k\right)$.

(ii) $\implies$ (i) Now the hypothesis is that the product exists. That is, there exists $w \in \mathbb{C}$ such that

$$\lim_{k\to\infty} \prod_{n=1}^{k} z_n = w.$$

As mentioned above, this means in particular that $z_n \to 1$. By working on the tail if needed, we may assume that $\log z_n$ exists for all $n$. Then, as log is continuous on a region removed from the negative half-line,

$$\sum_{k=1}^{n} \log z_k = \log\left(\prod_{k=1}^{n} z_k\right) \xrightarrow[n\to\infty]{} \log\left(\prod_{k=1}^{\infty} z_k\right) = \log w$$

and the series converges. $\square$

**3.11.3. Proposition.**

*If $\operatorname{Re} z_n > -1$ for all $n$, the following statements are equivalent:*

*(i)* $\sum\limits_n \log(1+z_n)$ *converges absolutely;*

*(ii)* $\sum\limits_n z_n$ *converges absolutely.*

*Proof.* Consider the Taylor series for $\log(1+z)$ around $z=0$:

$$\log(1+z) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{z^k}{k}, \qquad |z| < 1.$$

Then for any $|z| < 1$ we have

$$\left|1 - \frac{\log(1+z)}{z}\right| = \left|\sum_{k=2}^{\infty} (-1)^k \frac{z^{k-1}}{k}\right| \leq \frac{1}{2} \sum_{k=2}^{\infty} |z|^{k-1} = \frac{|z|}{2(1-|z|)}.$$

When in particular $|z| \leq \frac{1}{2}$, the above inequality becomes

$$\frac{1}{2}|z| \leq |\log(1+z)| \leq \frac{3}{2}|z|, \qquad |z| \leq \frac{1}{2}. \tag{3.32}$$

Now, if $\sum_n |\log(1+z_n)| < \infty$, then $\log(1+z_n) \to 0$ and so $z_n \to 0$. In particular, $|z_n| \leq \frac{1}{2}$ for all $n$ sufficiently large. Then (3.32) shows that $\sum_n |z_n| < \infty$ by comparison. And, conversely, if $\sum_n |z_n| < \infty$ then $|z_n| \leq \frac{1}{2}$ for sufficiently large $n$, and again comparison via (3.32) gives us that $\sum_n |\log(1+z_n)| < \infty$. □

In light of Proposition 3.11.3, we use it to define absolute convergence (note that convergence of the product of $|z_n|$ does not imply convergence of the product of $z_n$).

**3.11.4. Definition.**

*When* $\operatorname{Re} z_n > 0$ *for all sufficiently large* $n$*, the product* $\prod_n z_n$ *is said to* ***converge absolutely*** *if* $\sum_n |\log(1+z_n)| < \infty$.

Now we can consider products of holomorphic functions.

***3.11.5. Proposition.***

*Let* $V \subset \mathbb{C}$ *be open and* $\{f_n\}$ *a sequence of functions* $f_n : V \to \mathbb{C}$*, holomorphic and such that none of the* $f_n$ *is identically zero. Suppose that* $\sum_n (f_n(z)-1)$ *converges absolutely and uniformly on compact subsets of* $V$*. Then* $\prod_n f_n(z)$ *converges in* $V$ *(uniformly on compact sets) to a holomorphic function* $f$*. If* $f(w) = 0$*, then there exists at most a finite set* $\{n_1, \ldots, n_r\} \subset \mathbb{N}$ *such that* $f_{n_j}(w) = 0$*. The multiplicity of* $w$ *as a zero of* $f$ *is the sum of the multiplicities of* $w$ *as a zero of* $f_{n_1}, \ldots, f_{n_r}$*.*

*Proof.* Since $\sum_n |f_n(z) - 1| < \infty$, we have that $f_n(z) \to 1$ and so eventually $\operatorname{Re} f_n(z) > 0$. By Proposition 3.11.3 we have that $\sum_n |\log f_n(z)| < \infty$, and then by Proposition 3.11.2 $\prod_n f_n(z)$ converges (absolutely). As for the uniform convergence, given $K \subset V$ compact we have that $\sum_n (f_n(z) - 1)$ converges uniformly on $K$ by hypothesis. So there exists $n_0$ such that $|f_n(z) - 1| < \frac{1}{2}$ for all $z \in K$ and all $n \geq n_0$. Then $\operatorname{Re} f_n(z) > 0$ and $|\log f_n(z)| \leq \frac{3}{2}\,|f_n(z) - 1|$ for all $z \in K$ and all $n \geq n_0$ by (3.32). By comparison the series

$$g(z) = \sum_{n=n_0}^{\infty} \log f_n(z)$$

converges uniformly on $K$. The uniform convergence makes $g$ continuous, and hence bounded on $K$; so there exists $c > 0$ with $|\operatorname{Re} g(z)| \leq c$ for all $z \in K$. Fix $\varepsilon > 0$; by continuity of the exponential there exists $\delta > 0$ such that $|1 - e^z| < \varepsilon\, e^{-c}$ if $|z| < \delta$. Choose $m_0$ such that

$$\left| g(z) - \sum_{n=n_0}^{m} \log f_n(z) \right| < \delta, \qquad\qquad z \in K$$

for all $m \geq m_0$. For all such $m$,

$$\begin{aligned}
\left| \prod_{n=n_0}^{m} f_n(z) - e^{g(z)} \right| &= \left| \exp\Big( \sum_{n=n_0}^{m} \log f_n(z) \Big) - \exp(g(z)) \right| \\
&= |e^{g(z)}| \left| \exp\Big( \sum_{n=n_0}^{m} \log f_n(z) - g(z) \Big) - 1 \right| \\
&\leq e^{c}\, \varepsilon\, e^{-c} = \varepsilon.
\end{aligned}$$

Thus $\prod_n f_n(z)$ converges uniformly on $K$. We have

$$\prod_{n=1}^{\infty} f_n(z) = \Big( \prod_{n=1}^{n_0 - 1} f_n(z) \Big)\, e^{g(z)}.$$

As $e^{g(z)}$ is never zero, $f(w) = 0$ if and only if $f_n(w) = 0$ for some $n < n_0$. If $n_1, \ldots, n_r$ denote all the indices such that $f_{n_j}(w) = 0$, then there exists positive integers $m_1, \ldots, m_r$ such that $f_{n_j}(z) = (z - w)^{m_j}\, h_j(z)$, with $h_j(w) \neq 0$. Then the product can be written in the form

$$\prod_{n=1}^{\infty} f_n(z) = \prod_{j=1}^{r} (z - w)^{m_j}\, \tilde{h}(z),$$

where $\tilde{h}(w) \neq 0$. So the multiplicity of $w$ as a zero of the product is $m_1 + \cdots + m_r$. □

**3.11.6. Definition.**

*The functions*

$$E_p(z) = (1-z)\exp\left(\sum_{k=1}^{p} \frac{z^k}{k}\right), \qquad p = 0, 1, 2, \ldots,$$

*where the sum is interpreted as* 1 *when* $p = 0$, *are called* ***elementary factors***.

Since the exponential is never zero, the function $E_p(z/w)$ has a simple zero at $z = w$. The following estimate is crucial.

***3.11.7. Lemma.***

*Let* $p \in \mathbb{N} \cup \{0\}$ *and* $z \in \overline{\mathbb{D}}$. *Then*

$$|1 - E_p(z)| \le |z|^{p+1}. \tag{3.33}$$

*Proof.* When $p = 0$, (3.33) is an equality. So we fix $p \ge 1$. As $E_p$ is entire and $E_p(0) = 1$, its power series expansion at $z = 0$ is of the form

$$E_p(z) = 1 + \sum_{k=1}^{\infty} a_k z^k.$$

The derivative of $E_p$ can be easily calculated using the formula for the partial sums of the geometric series. Equating the derivatives on the two expressions for $E_p$,

$$-z^p \exp\left(\sum_{k=1}^{p} \frac{z^k}{k}\right) = \sum_{k=1}^{\infty} k\, a_k z^{k-1}.$$

As no monomial of degree $p-1$ or less can appear on the left hand side, we deduce that $a_1 = \cdots = a_p = 0$. A second observation is that the series expansion on the left has all negative coefficients; hence $a_k \le 0$ for all $k \ge p+1$. Thus

$$\sum_{k=p+1}^{\infty} |a_k| = -\sum_{k=p+1}^{\infty} a_k = 1 - E_p(1) = 1.$$

Now we can estimate, for $z \in \overline{\mathbb{D}}$,

$$\begin{aligned} |1 - E_p(z)| &= \left|\sum_{k=1}^{\infty} a_k\, z^k\right| \le \sum_{k=p+1}^{\infty} |a_k|\, |z|^k = |z|^{p+1} \sum_{k=p+1}^{\infty} |a_k|\, |z|^{k-p-1} \\ &\le |z|^{p+1} \sum_{k=p+1}^{\infty} |a_k| = |z|^{p+1}. \end{aligned}$$

□

We remark that the condition $\lim_n |z_n| = \infty$ in Proposition 3.11.8 allows only finitely many repetitions for any value taken by the $z_n$.

**3.11.8. Proposition.**

*Let $\{z_n\} \subset \mathbb{C}$ with $\lim_n |z_n| = \infty$. If $\{p_n\} \subset \{0\} \cup \mathbb{N}$ satisfies*

$$\sum_{n=1}^{\infty} \left(\frac{r}{|z_n|}\right)^{p_n+1} < \infty, \qquad r > 0, \tag{3.34}$$

*then the product*

$$f(z) = \prod_{n=1}^{\infty} E_{p_n}(z/z_n) \tag{3.35}$$

*converges for all $z$ and defines an entire function with zeros precisely the $\{z_n\}$. If $w$ appears in $\{z_n\}$ $m$ times, then $w$ is a zero of multiplicity $m$ for $f$. The condition (3.34) is satisfied in the particular case where $p_n = n - 1$.*

*Proof.* By Lemma 3.11.7

$$|1 - E_{p_n}(z/z_n)| \le \left|\frac{z}{z_n}\right|^{p_n+1} \le \left(\frac{r}{|z_n|}\right)^{p_n+1}, \qquad |z| \le r \le |z_n|.$$

Fix $r > 0$. By hypothesis there exists $\delta > 0$ with $r/|z_n| < 1 - \delta$ for all $n$ sufficiently large. Then the above inequality guarantees that

$$\sum_{n=1}^{\infty} |1 - E_{p_n}(z/z_n)| \le \sum_{n=1}^{\infty} \left(\frac{r}{|z_n|}\right)^{p_n+1} < \infty$$

uniformly for $|z| \le r$. Hence (3.35) converges by Proposition 3.11.5.

The condition on the multiplicity follows from the definition of the elementary factors, and the fact that each elementary factor admits a unique root.

If $p_n = n-1$, given $r > 0$ there exists $n_0$ such that $|z_n| > 2r$ for all $n \ge n_0$. Then the series in (3.34) is bounded by the geometric series $\sum_n 2^{-n}$. □

The factor $r^{p_n+1}$ in Proposition 3.11.8 can be ignored when the $p_n$ are bounded, as it does not affect convergence.

Now that we know how to construct entire functions with prescribed zeros, we are in position to characterize entire functions. We remark that the larger the $p_n$, the easier it is to satisfy (3.34); but one wants the smallest possible $p_n$, as this makes the powers in the exponential in $E_{p_n}$ smaller. We know from Corollary 3.6.3 that any holomorphic function has at most countably many zeros (Exercise 3.6.2).

**3.11.9. Theorem. (Weierstrass' Factorization)**

*Let $f$ be entire and $\{z_n\}$ the list of its zeros (with repetitions according to multiplicity). Let $m$ be the order of 0 as a zero of $f$ (so $m = 0$ means $f(0) \neq 0$). Then there exists $g$ entire and a sequence $\{p_n\} \subset \{0\} \cup \mathbb{N}$ such that*

$$f(z) = z^m \, e^{g(z)} \prod_{n=1}^{\infty} E_{p_n}\left(\frac{z}{z_n}\right).$$

*Proof.* By Proposition 3.11.8 there exist integers $p_n$ (explicitly, we can choose $p_n = n - 1$) such that the entire function

$$h(z) = z^m \prod_{n=1}^{\infty} E_{p_n}(z/z_n)$$

has precisely the same zeros as $f$, with the same multiplicities. Then $f/h$ has only removable singularities, so it is an entire function with no zeros. It follows that $(f/h)'/(f/h)$ is entire. Let $g = (f/h)'/(f/h)$. Then

$$\big((f/h)\, e^{-g}\big)' = \big(\,(f/h)' - g\,(f/h)\,\big)\, e^{-g} = 0.$$

Hence there exists $c \in \mathbb{C}$ such that $f/h = ce^g$, and so $f(z) = e^{g(z)}\, h(z)$ after replacing $g$ with $c_0 g$ for some $c_0 \in \mathbb{C}$ with $e^{c_0} = c$. □

With enough information about the zeros, sometimes one can also gather intel about the function $g$ in Theorem 3.11.9. Here is a concrete situation that will be needed in Section C.2.

**3.11.10. Proposition.**

*Let $f$ be entire with zeros $\{z_n\}$, and such that $f(0) = 1$ and*

*(i)* $\displaystyle\sum_{n=1}^{\infty} |z_n|^{-1} < \infty$;

*(ii) for each $\varepsilon > 0$, $|f(z)| \leq C_\varepsilon \exp(\varepsilon\,|z|)$ for some constant $C_\varepsilon > 0$.*

*Then*

$$f(z) = \prod_{n=1}^{\infty} (1 - z_n^{-1} z).$$

*Proof.* The condition on the zeros guarantees, via Proposition 3.11.8, that we can choose $p_n = 0$ for all $n$ in the proof of Theorem 3.11.9. Together with

$f(0) = 1$, this means that

$$f(z) = e^{h(z)} \prod_{n=1}^{\infty} (1 - z_n^{-1} z)$$

for some entire function $h$. Let $g(z) = \prod_{n=1}^{\infty} (1 - z_n^{-1} z)$. We want to show that $f = g$. Since both are entire (see Proposition 3.11.8) it is enough to show that they agree on a small disk around 0 (Corollary 3.6.2).

The assumption $f(0) = 1$ guarantees that $z_n \neq 0$ for all $n$. Fix $R > 1$ and let

$$K_R = \{n : \ |z_n| \leq R\}, \qquad J_R = \{n : \ |z_n| > R\}.$$

We have $K_R \cup J_R = \mathbb{N}$, and $K_R$ is finite. Because of this and that $f(0) = 1$ there exists $\delta > 0$ with $\delta < 1$ and such that $f_R(z) \in B_{1/2}(1)$, where

$$f_R(z) = f(z) \prod_{n \in K_R} (1 - z_n^{-1}\, z)^{-1}.$$

Then by working with $|z| < \delta$ we guarantee that $\log f_R$ is well-defined. We will use the principal branch of log, namely $-\pi \leq \theta < \pi$. We also define

$$\beta_R(z) = -\sum_{n \in J_R} \log(1 - z_n^{-1} z), \qquad\qquad |z| < \delta.$$

This makes sense because for $n \in J_R$ we have $|z_n^{-1} z| \leq \delta/R < \delta < 1$ and the logarithms are well-defined. The reason for these definitions is that

$$\begin{aligned} \beta_R(z) + \log f_R(z) &= -\sum_{n \in J_R} \log(1 - z_n^{-1} z) - \sum_{n \in K_R} \log(1 - z_n^{-1} z) + \log f(z) \\ &= -\sum_{n} \log(1 - z_n^{-1} z) + \log f(z) = \log f(z) - \log g(z). \end{aligned}$$

So if we show that $\beta_R$ and $\log f_R$ go to zero on $\delta\mathbb{D}$ as $R$ goes to infinity, we will have that $f = g$. We note that the logarithm applied to the infinite product $g$ does give the series of the logarithms, by continuity.

Since $\delta < 1$, there exists $c > 0$ such that $|\log(1 + z)| \leq c|z|$ for $|z| < \delta$ (one way to see this is that on that little ball, $l(z) = \log(1 + z)$ is analytic with $l(0) = 0$; another is to show that $\frac{1}{z} \log(1 + z))$ is bounded on the little ball, by looking at the real and imaginary parts). This gives us

$$|\beta_R(z)| \leq \sum_{n \in J_R} |\log(1 - z_n^{-1} z)| \leq c\, \delta \sum_{n \in J_R} |z_n|^{-1}.$$

As $R$ goes to infinity we have the tail of the convergent series $\sum_n |z_n|^{-1}$, and so $\beta_R(z) \to 0$.

As for $f_R$, because $K_R$ is finite the product in the formula is bounded. So there exists a constant $\alpha > 0$ such that $|f_R(z)| \leq \alpha\, |f(z)|$ for all $|z| < \delta$.

This gives us, for a fixed $\varepsilon > 0$,
$$|f_R(z)| \le \alpha\, C_\varepsilon \, \exp(\varepsilon\, |z|), \qquad\qquad |z| < \delta.$$
Then
$$\operatorname{Re} \log f_R(z) = \log |f_R(z)| \le \log(\alpha + C_\varepsilon) + \varepsilon\, \delta.$$
By Corollary 3.9.6, since $\log f_R(0) = 1$, and using the Maximum Modulus Principle (Corollary 3.7.3)
$$\max_{|z|\le\delta} |\log f_R(z)| \le \max |z| = \delta |\log f_R(z)| \le \frac{2\delta}{R-\delta}\big(\log(\alpha + C_\varepsilon) + \varepsilon\, \delta\big).$$
It is then clear that $\lim_{R\to\infty} \log f_R(z) = 0$ for all $z \in \delta\mathbb{D}$, and so $f(z) = g(z)$ on $\delta\mathbb{D}$ and hence everywhere. □

*3.11.11. Remark*. Another technique one can use to work out the function $g$ above is to use the **logarithmic derivative**. This is, for a function $f$, the function $f'/f$. This function "behaves like the logarithm" in the sense that it transforms products into sums. We refer to Exercise 3.11.1.

The following is an improvement over Theorem 3.11.9 when the growth of the function is known.

***3.11.12. Theorem. (Hadamard Factorization)***

*Let $f$ be entire and $\{z_n\}$ the list of its zeros (with repetitions according to multiplicity). Let $m$ be the order of $0$ as a zero of $f$ (so $m = 0$ means $f(0) \neq 0$). Suppose also that there exists $s_0 > 0$ such that there exist $a, b > 0$ such that $|f(z)| \le a e^{b|z|^s}$ for every $s > s_0$. Then there exists a polynomial $p$ with $\deg p \le k = \lfloor s_0 \rfloor$ such that*
$$f(z) = z^m\, e^{p(z)} \prod_{n=1}^{\infty} E_k\left(\frac{z}{z_n}\right).$$

*Proof.* Exercises 3.11.7 to 3.11.12. □

### 3.11.1. Exercises.

(3.11.1) Let $h, k : V \to \mathbb{C}$ be functions defined on a region $V$. Show that if both functions are nonzero and differentiable at $z = z_0$, then
$$\frac{(hk)'(z_0)}{h(z_0)k(z_0)} = \frac{h'(z_0)}{h(z_0)} + \frac{k'(z_0)}{k(z_0)}$$
and
$$\frac{(h/k)'(z_0)}{h(z_0)/k(z_0)} = \frac{h'(z_0)}{h(z_0)} - \frac{k'(z_0)}{k(z_0)}.$$

(3.11.2) Find the Weierstrass factorization of the entire function $f(z) = \sin \pi z$.

(3.11.3) Show that Theorem 3.11.9 holds for entire functions with the convention that the empty product is equal to 1. That is, given $f$ entire with no zeros, there exists $g$ entire with $f = e^g$.

(3.11.4) Show that the Weierstrass product

$$f(z) = z \prod_{n=1}^{\infty} \left(1 - \frac{z}{n}\right) e^{z/n}$$

converges uniformly on compact sets and defines an entire function.

(3.11.5) Find an entire function with simple zeros at $z = n^2$ for all $n \in \mathbb{N}$. Ensure that you choose the minimal $k$ in each factor $E_k(z)$.

(3.11.6) Considering the Gamma function as a meromorphic function like in Exercise 3.10.20, show that $f(z) = 1/\Gamma(z)$ is entire and

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n}$$

where $\gamma$ is the **Euler constant**

$$\gamma = \sum_{n=1}^{\infty} \frac{1}{n} - \log\left(1 + \frac{1}{n}\right).$$

*Exercises 3.11.7-3.11.12 make up the proof of Theorem 3.11.12. They are interesting and not trivial.*

(3.11.7) Show the inequalities

$$|E_n(z)| \geq e^{-2|z|^{n+1}}, \qquad |z| < \frac{1}{2}$$

and

$$|E_n(z)| \geq |1 - z| e^{c_n |z|^n}, \qquad |z| \geq \frac{1}{2}$$

for constants $c_n > 0$, $n \in \mathbb{N}$.

(3.11.8) Show that if $\{z_n\}$ is a sequence with $\sum_n |z_n|^{-s} < \infty$ for any $s > k$, such that $|\{|z_n| < r\}| \leq c_0 r^s$, and $|z - z_n| > |z_n|^{-k-1}$ for all $n$, then there exist $a, b > 0$ such that if $s \in (k, k+1)$

$$\left| \prod_{n=1}^{\infty} E_k\left(\frac{z}{z_n}\right) \right| \geq ae^{-b|z|^s}.$$

(3.11.9) Let $R > 0$ and $f$ analytic on $R\overline{\mathbb{D}}$, with $f(0) \neq 0$ and $f \neq 0$ on the boundary. Denote the finitely many roots of $f$ by $\{z_1, \ldots, z_m\}$, counting multiplicities. Show Jensen's formula

$$\log |f(0)| = \sum_n \log \frac{|z_n|}{R} + \frac{1}{2\pi} \int_0^{2\pi} \log |f(Re^{i\theta})| \, d\theta.$$

by going through the following steps.

(i) Show that it is enough to prove the case $R = 1$.

(ii) Prove the case where $f$ has not roots.

(iii) Show that the formula behaves nicely with products.

(iv) Prove the case where $f(z) = z - z_1$.

(v) Conclude the general case.

(3.11.10) Let $R > 0$ and $f$ analytic on $R\overline{\mathbb{D}}$, with $f(0) \neq 0$ and $f \neq 0$ on the boundary. Denote the finitely many roots of $f$ by $\{z_1, \ldots, z_m\}$, counting multiplicities. Let $\mathfrak{n}(r)$ denote the number of nonzero roots of $f$ in $R\mathbb{D}$. Show the following variation of Jensen's formula:

$$\log|f(0)| = -\int_0^R \frac{\mathfrak{n}(x)}{x}\,dx + \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|\,d\theta. \tag{3.36}$$

(3.11.11) Let $f$ be entire with $\{z_n\}$ its nonzero roots and such that $|f(z)| \leq ae^{b|z|^s}$ for certain constants $a, b$ and all $s > s_0 > 0$.

(i) Show that there exists $c > 0$ with $\mathfrak{n}(r) \leq cr^{s_0}$ for large enough $r$.

(ii) Show that

$$\sum_{n=1}^{\infty} \frac{1}{|z_n|^s} < \infty$$

for all $s > s_0$.

(3.11.12) Use Exercises 3.9.2, 3.11.7, 3.11.8 and 3.11.11 to write the proof of Theorem 3.11.12.

(3.11.13) Consider the entire function $s(z) = \displaystyle\sum_{n=0}^{\infty} \frac{z^n}{(2n+1)!}$.

(i) Show that

$$s(z^2) = \frac{\sinh z}{z}.$$

(ii) Show that

$$s(z) = \prod_{n=1}^{\infty}\left(1 + \frac{z}{n^2\pi^2}\right).$$

CHAPTER

4

# Hilbert spaces

Hilbert spaces are certain vector spaces where the scalar field is either $\mathbb{R}$ or $\mathbb{C}$. We will see along the way that there are several reasons to have a preference for the complex case, so with few exceptions our field of choice throughout the rest of the book will be $\mathbb{C}$. We recall the notation $\mathbb{D} = \{\lambda \in \mathbb{C} : \ |\lambda| < 1\}$ and $\mathbb{T} = \{\lambda \in \mathbb{C} : \ |\lambda| = 1\}$ for the **unit disk** and **unit circle**, respectively. Sometimes, when emphasizing the view of $\mathbb{C}$ as a metric space, we may use $B_1(0)$ instead of $\mathbb{D}$. We also use $B_r(x)$ to denote, in any metric space, the ball of radius $r$ centered at $x$.

The Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ carry a natural norm,

$$\|x\| = \left(\sum_{j=1}^n |x_j|^2\right)^{1/2},$$

that allows us to write the distance between $x$ and $y$ as $\|x - y\|$; moreover, this norm is given by an *inner product*. So, besides being vector spaces, we may see these spaces as metric spaces—allowing us to do *analysis*, i.e. limits, continuity. A common disconnect when going through algebra and analysis is that, from Calculus, one is familiar with continuous functions on $\mathbb{R}^n$; and, from Linear Algebra, one is familiar with linear maps in $\mathbb{R}^n$. But the question of continuity of these maps does not usually arise. This question is itself trivial, but we will see that it becomes non-trivial when going into infinite dimension.

Furthermore, the inner product gives us a notion of *orthogonality*, that allows us to think geometrically; on the technical side, it allows us to choose certain bases with very nice properties. This situation is common enough that it deserves its own study, which becomes richer in the infinite-dimensional case. Hence, Hilbert Spaces.

## 4.1. Basic Definitions

We begin by defining what a Hilbert space is. This requires us to define the notion of inner product, and we take the opportunity to define related terminology.

**4.1.1. Definition.**

*A (complex)* ***Hilbert space*** *is a triple* $(\mathcal{H}, +, \langle\,,\,\rangle)$ *where* $\mathcal{H}$ *is a set,* $+$ *is a binary operation that makes* $\mathcal{H}$ *into a vector space over* $\mathbb{C}$*, and* $\langle\,,\,\rangle$ *is an* ***inner product****, i.e. a function* $\mathcal{H} \times \mathcal{H} \to \mathbb{C}$ *such that*

(i) $\langle a\xi + \eta, \nu\rangle = a\,\langle \xi, \nu\rangle + \langle \eta, \nu\rangle$ *for all* $\xi, \eta, \nu \in \mathcal{H}$*,* $a \in \mathbb{C}$*;*

(ii) $\langle \xi, b\eta + \nu\rangle = \bar{b}\,\langle \xi, \eta\rangle + \langle \xi, \nu\rangle$ *for all* $\xi, \eta, \nu \in \mathcal{H}$*,* $b \in \mathbb{C}$*;*

(iii) $\langle \xi, \eta\rangle = \overline{\langle \eta, \xi\rangle}$ *for all* $\xi, \eta \in \mathcal{H}$*;*

(iv) $\langle \xi, \xi\rangle \geq 0$ *for all* $\xi \in \mathcal{H}$*;*

(v) $\langle \xi, \xi\rangle = 0$ *if and only if* $\xi = 0$*;*

(vi) $\mathcal{H}$ *is complete with the topology given by the distance* $d(\xi, \eta) = \langle \xi - \eta, \xi - \eta\rangle^{1/2}$.

*A function satisfying (i)–(ii) is often called a* ***sesquilinear form****; it is* ***hermitian*** *when it satisfies (iii), and* ***positive*** *when it also satisfies (iv). When (vi) is not satisfied, we say that* $\mathcal{H}$ *is a* ***pre-Hilbert space****, since its completion will be a Hilbert space.*

One can also define a **real Hilbert space** by just replacing complex with real above. In the real case, "sesquilinear" is reduced to "bilinear". We note that combining (i) and (iii) we immediately get (ii), but we include it in the definition to be able to define the term sesquilinear.

*4.1.2. Examples* . Here are some of the most common examples of Hilbert spaces.

(a) For any $n \in \mathbb{N}$, $\mathbb{C}^n$ is a Hilbert space with the entrywise sum and canonical inner product, i.e. $\langle x, y\rangle = y^*x$, where here $y^*$ is the conjugate transpose of $y$. We will see later that—up to isomorphism—this is the full list of finite-dimensional (complex) Hilbert spaces.

(b) The set

$$\ell^2(\mathbb{N}) = \{x : \mathbb{N} \to \mathbb{C} : \sum_n |x(n)|^2 < \infty\}$$

is a Hilbert space when considered with pointwise addition and inner product given by $\langle x, y\rangle = \sum_n x(n)\overline{y(n)}$. This is the straightforward infinite-dimensional generalization of the $\mathbb{C}^n$ example above. One can check that the series for $\langle x, y\rangle$ always converges for $x, y \in \ell^2(\mathbb{N})$, by using Hölder's Inequality.

(c) More generally, for any measure space $(X, \Sigma, \mu)$, the set

$$L^2(X, \Sigma, \mu) = \{[f] : \ f : X \to \mathbb{C} \text{ is measurable, with } \int |f|^2 \, d\mu < \infty\},$$

where $[f] = \{g : \ f - g = 0 \ \text{ a.e.}\}$ is the equivalence class of $f$ up to a nullset, is a Hilbert space when considered with pointwise addition of representatives and the inner product

$$\langle f, g\rangle = \int f\bar{g}\, d\mu.$$

(d) The vector space $M_n(\mathbb{C})$ of complex $n \times n$ matrices, with the inner product

$$\langle X, Y\rangle = \mathrm{Tr}(Y^* X),$$

is a Hilbert space.

It is not immediately obvious that the spaces in (b) and (c) are complete, but that is precisely the contents of the Riesz–Fischer Theorem (2.8.12). Let us also remark that the examples in (a) and (b) are particular cases of (c) (when using the counting measure).

Another feature of infinite-dimensional Hilbert spaces is the existence of proper dense subspaces. For instance,

$$c_{00} = \{x : \mathbb{N} \to \mathbb{C} : \ x \text{ is nonzero on a finite set}\}$$

is a dense subspace in $\ell^2(\mathbb{N})$. And $C[0, 1]$ is a dense subspace of $L^2[0, 1]$ as are the subspace of polynomials and the subspace of simple functions.

**4.1.1. Exercises.**

(4.1.1) Prove that each of the spaces in Examples 4.1.2 is actually a Hilbert space.

(4.1.2) Show that $c_{00}$ is dense in $\ell^2(\mathbb{N})$.

(4.1.3) Find examples of pre-Hilbert spaces which are not Hilbert spaces.

## 4.2. The role of the inner product in the topological structure

The inner product on a Hilbert space plays two main roles: first, it induces a norm, and thus a topology; and second, it provides a notion of orthogonality.

**4.2.1. Definition.**

*A **norm** on a vector space $\mathcal{H}$ is a function $\|\cdot\| : \mathcal{H} \to [0,\infty)$ such that*

*(i)* $\|\xi\| = 0$ *implies* $\xi = 0$;

*(ii)* *(**Homogeneity**)* $\|b\xi\| = |b|\,\|\xi\|$ *for all* $b \in \mathbb{C}$, $\xi \in \mathcal{H}$;

*(iii)* *(**Triangle Inequality**)* $\|\xi+\eta\| \leq \|\xi\| + \|\eta\|$, *for all* $\xi, \eta \in \mathcal{H}$.

Given a pre-Hilbert space $\mathcal{H}$, its inner product induces a norm by

$$\|\xi\| = \langle \xi, \xi\rangle^{1/2}, \qquad \xi \in \mathcal{H}. \tag{4.1}$$

It is not patently obvious that (4.1) satisfies the triangle inequality—(iii) in Definition 4.2.1—but we will verify it very soon (Corollary 4.2.3). For this we will need the Cauchy–Schwarz inequality, to be proven below.

A straightforward computation shows that, for any $\xi, \eta \in \mathcal{H}$,

$$\|\xi+\eta\|^2 + \|\xi-\eta\|^2 = 2\|\xi\|^2 + 2\|\eta\|^2. \tag{4.2}$$

This equality is called the **Parallelogram Identity**. It is a clearly necessary (but also sufficient!) condition for a norm to be induced by an inner product (Theorem 4.2.7). For instance the norm, on $\mathbb{C}^n$, given by

$$\|x\|_1 = \sum_{k=1}^n |x_k|$$

does not satisfy the parallelogram identity, and thus cannot proceed from an inner product: we have

$$\|(1,0)+(0,1)\|_1^2 + \|(1,0)-(0,1)\|_1^2 = 8 \neq 4 = 2\|(1,0)\|_1^2 + 2\|(0,1)\|_1^2.$$

Here is a straightforward but ever-present equality, that will be used often and without mention:

$$\|\xi+\eta\|^2 = \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\,\langle \xi, \eta\rangle. \tag{4.3}$$

The next result, the **Cauchy–Schwarz Inequality**, is an essential tool. The formulation in terms of a “sesquilinear and positive” form is due to the fact that one does not need the implication "$\phi(\xi,\xi)=0 \implies \xi=0$", so the inequality applies to forms that are not inner products.

**4.2.2. Theorem. (Cauchy–Schwarz inequality)**

*Let $\mathcal{H}$ be a vector space and $\varphi:\mathcal{H}\times\mathcal{H}\to\mathbb{C}$ be sesquilinear and positive. Then*

$$|\varphi(\xi,\eta)| \le \varphi(\xi,\xi)^{1/2}\,\varphi(\eta,\eta)^{1/2}, \qquad \xi,\eta\in\mathcal{H}. \tag{4.4}$$

*If $\varphi$ is an inner product, equality in (4.4) occurs if and only if there exists $c\in\mathbb{C}$ with $\eta=c\xi$ or $\xi=c\eta$.*

*Proof.* Fix $\xi,\eta\in\mathcal{H}$ and let $t\in\mathbb{C}$. Then

$$0\le\varphi(\xi+t\eta,\xi+t\eta)=\varphi(\xi,\xi)+|t|^2\varphi(\eta,\eta)+2\mathrm{Re}\,t\varphi(\eta,\xi). \tag{4.5}$$

Let $\varphi(\eta,\xi)=e^{i\theta}|\varphi(\eta,\xi)|$ be the polar form of $\varphi(\eta,\xi)$. Next we consider the inequality (4.5) for the same $\xi,\eta$ and the complex number $e^{-i\theta}t$, with $t\in\mathbb{R}$. So we have

$$0\le\varphi(\xi+e^{-i\theta}t\eta,\xi+e^{-i\theta}t\eta)=\varphi(\xi,\xi)+t^2\varphi(\eta,\eta)+2t|\varphi(\eta,\xi)|.$$

We have a real quadratic expression on $t$, and for it to be always nonnegative we need

$$4|\varphi(\eta,\xi)|^2\le 4\varphi(\xi,\xi)\varphi(\eta,\eta).$$

Dividing by 4 and taking square root—all quantities are non-negative—yields (4.4).

For the case of equality in (4.4) when $\varphi$ is an inner product, assume first that $\eta\ne0$. Then, writing $\bar\varphi(\xi)=\varphi(\xi,\xi)$ to shorten the notation a bit,

$$\begin{aligned}
\varphi(\xi,\xi)&=\bar\varphi\left(\xi-\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta+\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)\\
&=\bar\varphi\left(\xi-\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)+\bar\varphi\left(\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)+2\mathrm{Re}\,\varphi\left(\xi-\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta,\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)\\
&=\bar\varphi\left(\xi-\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)+\bar\varphi\left(\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)=\bar\varphi\left(\xi-\frac{\varphi(\xi,\eta)}{\varphi(\eta,\eta)}\,\eta\right)+\varphi(\xi,\xi),
\end{aligned}$$

where the hypothesis $|\varphi(\xi,\eta)|^2=\varphi(\xi,\xi)\varphi(\eta,\eta)$ was used in the last equality. Since we are assuming, for the case of equality, that $\varphi$ is an inner product, we conclude that $\xi-\varphi(\xi,\eta)/\varphi(\eta,\eta)\,\eta=0$. If $\eta=0$ and $\xi\ne0$, we can repeat the proof with the roles reversed. ☐

In the particular case where $\varphi$ is the inner product of a Hilbert space $\mathcal{H}$, we may write the Cauchy–Schwarz inequality as

$$|\langle\xi,\eta\rangle|\le\|\xi\|\,\|\eta\|, \qquad \xi,\eta\in\mathcal{H}.$$

Also, when $\varphi$ is an inner product, one can have this slightly more direct proof of the Cauchy–Schwarz inequality: when $\eta \neq 0$ and so $\varphi(\eta,\eta) > 0$, we have

$$\begin{aligned}
0 &\leq \varphi\Big( \varphi(\eta,\eta)^{1/2}\,\xi - \frac{\overline{\varphi(\xi,\eta)}}{\varphi(\eta,\eta)^{1/2}}\,\eta, \varphi(\eta,\eta)^{1/2}\,\xi - \frac{\overline{\varphi(\xi,\eta)}}{\varphi(\eta,\eta)^{1/2}}\,\eta\Big) \\
&= \varphi(\eta,\eta)\varphi(\xi,\xi) + |\varphi(\xi,\eta)|^2 - 2\mathrm{Re}\,|\varphi(\xi,\eta)|^2 \\
&= \varphi(\eta,\eta)\varphi(\xi,\xi) - |\varphi(\xi,\eta)|^2,
\end{aligned}$$

which gives the inequality. But we cannot use this for a general sesquilinear and positive $\varphi$, because it does not deal with the case $\varphi(\eta,\eta) = 0$; until we have Cauchy–Schwarz, we have no way to say that if $\varphi(\eta,\eta) = 0$ then $\varphi(\xi,\eta) = 0$ for any $\xi$.

***4.2.3. Corollary. (Triangle Inequality)***

*Let $\mathcal{H}$ be a pre-Hilbert space, $\xi,\eta \in \mathcal{H}$. Then*

$$\|\xi+\eta\| \leq \|\xi\| + \|\eta\|,$$

*with equality if and only if there exists $c \geq 0$ with $\eta = c\xi$ or $\xi = c\eta$.*

*Proof.* We have

$$\begin{aligned}
\|\xi+\eta\|^2 = \langle \xi+\eta, \xi+\eta\rangle &= \|\xi\|^2 + \|\eta\|^2 + 2\mathrm{Re}\,\langle \xi,\eta\rangle \\
&\leq \|\xi\|^2 + \|\eta\|^2 + 2|\langle \xi,\eta\rangle| \\
&\leq \|\xi\|^2 + \|\eta\|^2 + 2\|\xi\|\,\|\eta\| = (\|\xi\| + \|\eta\|)^2.
\end{aligned}$$

Equality in the triangle inequality forces the inequality above to be an equality, and so we get equality in Cauchy–Schwarz, which gives the desired relation between $\xi$ and $\eta$. The reason that $c \geq 0$ is that $\|\xi + c\xi\| = |1+c|\,\|\xi\|$ and $\|\xi\| + \|c\xi\| = (1+|c|)\|\xi\|$. So as long as $\xi \neq 0$, $|1+c| = 1+|c|$, which forces $c \geq 0$. ☐

Although trivial in appearance, this next result has practical and theoretical consequences that we will exploit further down the line.

***4.2.4. Proposition.***

*If $\mathcal{H}$ is a pre-Hilbert space then, for any $\xi \in \mathcal{H}$,*

$$\|\xi\| = \sup\{|\langle \xi,\eta\rangle| : \ \eta \in \mathcal{H},\ \|\eta\| = 1\}.$$

*Proof.* Since $|\langle\xi,\eta\rangle| \leq \|\xi\|\,\|\eta\| = \|\xi\|$, it is clear that the right-hand-side is always less than or equal the left-hand-side. But $\|\xi\| = \langle\xi, \xi/\|\xi\|\rangle$, which shows the equality. $\square$

*4.2.5. Remark* . The result in Proposition 4.2.4 is a particular case of the well-known measure-theoretic result (Proposition 5.6.8) that, for any $f \in L^p$,

$$\|f\|_p = \sup\left\{\int_X |fg| : \ \|g\|_q = 1\right\},$$

where $p \geq 1$ and $p^{-1} + q^{-1} = 1$.

***4.2.6. Proposition. (Polarization Identity)***

*Let $\mathcal{H}$ be a vector space, $\xi, \eta \in \mathcal{H}$, and $\psi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ any sesquilinear form. Then*

$$\psi(\xi,\eta) = \frac{1}{4}\sum_{k=0}^{3} i^k \psi(\xi + i^k\eta, \xi + i^k\eta) \tag{4.6}$$

*Proof.* The straightforward proof is left to the reader (Exercise 4.2.3). $\square$

The Polarization Identity might look weird at first, but it will be used again and again. In the current context, it allows us to recover the inner product out of the norm:

***4.2.7. Theorem.***

*Let $\mathcal{H}$ be a normed space. The norm in $\mathcal{H}$ comes from an inner product if and only if it satisfies the parallelogram identity (4.2).*

*Proof.* The reader is asked to prove the Parallelogram Identity in Exercise 4.2.4.

Now suppose that we have a norm satisfying the Parallelogram Identity. Define

$$\langle\xi,\eta\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|\xi + i^k\eta\|^2.$$

Then

$$\langle\xi,\xi\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|\xi + i^k\xi\|^2 = \frac{1}{4}\sum_{k=0}^{3} i^k |1 + i^k|^2 \|\xi\|^2$$

$$= \frac{1}{4}\left(4\|\xi\|^2 + 2i\|\xi\|^2 - 0 - 2i\|\xi\|^2\right) = \|\xi\|^2 \geq 0,$$

giving us positivity, that $\|\xi\|^2 = \langle \xi, \xi\rangle$, and also the fact that if $\langle \xi, \xi\rangle = 0$, then $\xi = 0$. Also,

$$\langle \eta, \xi\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\|\eta + i^k\xi\|^2 = \frac{1}{4}\sum_{k=0}^{3} i^k\|\xi + i^{-k}\eta\|^2$$

$$= \overline{\frac{1}{4}\sum_{k=0}^{3} i^{-k}\|\xi + i^{-k}\eta\|^2} = \overline{\langle \xi, \eta\rangle}.$$

Next, let us prove that the product we defined is additive on both components. Note that we haven't used the Parallelogram Identity yet. Since $\langle \eta, \xi\rangle = \overline{\langle \xi, \eta\rangle}$, it is enough to show additivity on one side. We prove first that $\langle 2\xi, \eta\rangle = 2\langle \xi, \eta\rangle$. Indeed, using the Parallelogram Identity,

$$\begin{aligned}\|2\xi + i^k\eta\|^2 &= \|\xi + i^k\eta + \xi\|^2 \\ &= 2\|\xi + i^k\eta\|^2 + 2\|\xi\|^2 - \|i^k\eta\|^2.\end{aligned}$$

Adding from $k = 0$ to $k = 3$, and using that $\sum_{k=0}^{3} i^k = 0$ and $\|i^k\eta\| = \|\eta\|$, we obtain

$$\langle 2\xi, \eta\rangle = 2\langle \xi, \eta\rangle.$$

Also,

$$\langle \xi, 2\eta\rangle = \overline{\langle 2\eta, \xi\rangle} = 2\overline{\langle \eta, \xi\rangle} = 2\langle \xi, \eta\rangle.$$

Then

$$\begin{aligned}\langle \xi + \nu, \eta\rangle &= \frac{1}{4}\sum_{k=0}^{3} i^k\|\xi + \nu + i^k\eta\|^2 = \frac{1}{4}\sum_{k=0}^{3} i^k\left\|\xi + \frac{i^k\eta}{2} + \nu + \frac{i^k\eta}{2}\right\|^2 \\ &= \frac{1}{4}\sum_{k=0}^{3} i^k\left(2\left\|\xi + \frac{i^k\eta}{2}\right\|^2 + 2\left\|\nu + \frac{i^k\eta}{2}\right\|^2 - \|\xi - \nu\|^2\right) \\ &= \frac{1}{4}\sum_{k=0}^{3} i^k\left(2\left\|\xi + \frac{i^k\eta}{2}\right\|^2 + 2\left\|\nu + \frac{i^k\eta}{2}\right\|^2\right) \\ &= 2\langle \xi, \eta/2\rangle + 2\langle \nu, \eta/2\rangle = \langle \xi, \eta\rangle + \langle \nu, \eta\rangle.\end{aligned}$$

Thus the product is additive in both components. Note that

$$\langle 0, \eta\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\|i^k\eta\|^2 = 0.$$

Looking at scalar multiplication, from the additivity $\langle n\xi, \eta\rangle = n\langle \xi, \eta\rangle$ for all $n \in \mathbb{N}$. Also, from $\langle -\xi, \eta\rangle + \langle \xi, \eta\rangle = \langle -\xi + \xi, \eta\rangle = 0$, we get $\langle -\xi, \eta\rangle =$

$-\langle \xi, \eta \rangle$. Then, when $-n \in \mathbb{N}$,

$$\langle n\xi, \eta \rangle = -\langle -n\xi, \eta \rangle = -(-n)\langle \xi, \eta \rangle = n\langle \xi, \eta \rangle,$$

so the equality $\langle n\xi, \eta \rangle = n\langle \xi, \eta \rangle$ holds for all $n \in \mathbb{Z}$. Now $m\langle (\xi/m), \eta \rangle = \langle m(\xi/m), \eta \rangle = \langle \xi, \eta \rangle$, so

$$\langle \xi/m, \eta \rangle = \frac{1}{m} \langle \xi, \eta \rangle, \qquad m \in \mathbb{Z} \setminus \{0\}.$$

Combining the above,

$$\langle q\xi, \eta \rangle = q \langle \xi, \eta \rangle, \qquad q \in \mathbb{Q}.$$

For $r \in \mathbb{R}$, take a sequence $\{q_n\} \subset \mathbb{Q}$ with $q_n \to r$. We have, via the reverse triangle inequality,

$$\big| \|q_n\xi + i^k\eta\| - \|r\xi + i^k\eta\| \big| \leq \|q_n\xi - r\xi\| = |q_n - r| \, \|\xi\| \to 0,$$

and so

$$\begin{aligned} \langle r\xi, \eta \rangle &= \frac{1}{4} \sum_{k=0}^{3} i^k \|r\xi + i^k\eta\|^2 = \lim_n \frac{1}{4} \sum_{k=0}^{3} i^k \|q_n\xi + i^k\eta\|^2 = \lim_n \langle q_n\xi, \eta \rangle \\ &= \lim_n q_n \langle \xi, \eta \rangle = r\langle \xi, \eta \rangle. \end{aligned}$$

We have now scalar multiplication by the reals. Also,

$$i \langle \xi, \eta \rangle = \frac{1}{4} \sum_{k=0}^{3} i^{k+1} \|\xi + i^k\eta\|^2 = \frac{1}{4} \sum_{k=0}^{3} i^{k+1} \|i\xi + i^{k+1}\eta\|^2 = \langle i\xi, \eta \rangle.$$

As we already have additivity, the above results combine to give us $\langle a\xi, \eta \rangle = a\langle \xi, \eta \rangle$ for all $a \in \mathbb{C}$. Thus $\langle \cdot, \cdot \rangle$ defines an inner product. ☐

*4.2.8. Remark*. While so far we have used the polarization identity to characterize those normed spaces that are inner product spaces, its usefulness goes way beyond that, as it is a great tool in many cases where some information is known about the numbers $\psi(\xi, \xi)$ and the identity helps to extend the information to all $\psi(\xi, \eta)$. The reason why Proposition 4.2.6 is written in terms of a general sesquilinear form instead of the inner product $\langle \cdot, \cdot \rangle$ is that we will use it for many other forms, usually not generated by the inner product (since, most importantly, positivity is not required!). As an easy example of these kind of applications, suppose that $T : \mathcal{H} \to \mathcal{H}$ is a linear operator, and we know that $\langle T\xi, \xi \rangle = 0$ for all $\xi \in \mathcal{H}$. Then we apply Proposition 4.2.6 to the form $\psi(\xi, \eta) = \langle T\xi, \eta \rangle$ to conclude that $\langle T\xi, \eta \rangle = 0$ for all $\xi, \eta \in \mathcal{H}$. Taking $\eta = T\xi$ this implies that $T\xi = 0$ for all $\xi \in \mathcal{H}$.

*4.2.9. Remark*. The proof of Theorem 4.2.7 shows a bit more than what the theorem says, in the sense that it does not require the norm $\| \cdot \|$ to actually be a norm, since the triangle inequality plays no—direct—role. The only role of the triangle inequality in the proof of Theorem 4.2.7 is to show that $\|\cdot\|$ is continuous. In other words, any map $\gamma : \mathcal{H} \to [0, \infty)$ which is homogeneous and continuous with respect to multiplication by scalars (that, is the function $t \longmapsto \|t\xi + \eta\|$ is continuous), and that satisfies the parallelogram identity, can be used to define a positive sesquilinear form.

### 4.2.1. Exercises.

(4.2.1) Prove (4.3).

(4.2.2) Let $n \in \mathbb{N}$ and $\mathcal{H} = \mathbb{C}^n$. Show that $\mathcal{H}$ admits infinitely many inner products that are not multiples of each other, and that they all induce the same topology (for this, show that the induced norms are equivalent).

(4.2.3) Prove the Polarization Identity (4.6).

(4.2.4) Prove the Parallelogram Identity (4.2).

(4.2.5) Use the Parallelogram Identity (4.2) to show that none of the examples on Section 5.1 is a Hilbert space, with the exception of Example 5.1.4 and the case $p = 2$ in Examples 5.1.7 and 5.1.8.

(4.2.6) Fix $n \geq 3$, and let $\omega \in \mathbb{C}$ be an $n^{\text{th}}$ primitive root of unity. Let $\psi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form. Show that

$$\psi(\xi, \eta) = \frac{1}{n} \sum_{k=0}^{n-1} w^k \psi(\xi + \omega^k \eta, \xi + \omega^k \eta). \tag{4.7}$$

(4.2.7) Does (4.7) work for $n = 2$? Why?

(4.2.8) Let $\psi : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form. Show that

$$\psi(\xi, \eta) = \frac{1}{2\pi} \int_0^{2\pi} e^{it} \psi(\xi + e^{it}\eta, \xi + e^{it}\eta)\, dt. \tag{4.8}$$

(4.2.9) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta, \nu \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|\eta\| \leq 1$, $\|\nu\| \leq 1$. Fix $\varepsilon > 0$. Show that

$$\left\|\xi - \frac{\eta + \nu}{2}\right\| < \varepsilon \text{ implies } \|\xi - \eta\| < 2\sqrt{\varepsilon} \text{ and } \|\xi - \nu\| < 2\sqrt{\varepsilon}.$$

## 4.3. Orthogonality

From now on, $\mathcal{H}$ will denote a Hilbert space unless specifically mentioned. With the usual inner product in $\mathbb{R}^n$, one can check geometrically and algebraically that perpendicularity of two vectors is characterized by their inner product being equal to zero (since $\langle x, y\rangle = \|x\|\,\|y\|\cos\theta$, with $\theta$ the angle between $x$ and $y$). Having an arbitrary inner product allows us to define perpendicularity (we will say orthogonality) just algebraically.

**4.3.1. Definition.**

*Let $\mathcal{H}$ be an inner product space. Two vectors $\xi, \eta \in \mathcal{H}$ are said to be* ***orthogonal****, denoted $\xi \perp \eta$, if $\langle \xi, \eta\rangle = 0$. Similarly, we say that two subsets $A, B \subset \mathcal{H}$ are orthogonal, if $\xi \perp \eta$ whenever $\xi \in A$, $\eta \in B$.*

*The* ***orthogonal complement*** *of $A \subset \mathcal{H}$ is the set*

$$A^\perp = \{\xi \in \mathcal{H} : \ \xi \perp \eta \ \forall \eta \in A\}.$$

*4.3.2. Examples.* If $\mathcal{H} = \mathbb{C}^2$ with the usual inner product $\langle (a, b), (c, d)\rangle = a\,\bar{b} + c\,\bar{d}$, then

$$\{(a, b)\}^\perp = \{(c, d) : \ a\,\bar{b} + c\,\bar{d} = 0\} = \{z(-\bar{b}, \bar{a}) : \ z \in \mathbb{C}\}.$$

The second expression emphasizes the fact that in this case the orthogonal is one-dimensional.

Let $\mathcal{H} = \ell^2(\mathbb{N})$ and take $A = \{x \in \mathcal{H} : \ x(1) = x(2) = 0\}$. Then

$$A^\perp = \{x \in \mathcal{H} : x(n) = 0, \ n \geq 3\}.$$

It is easy to come up with other simple examples, but in general one should not expect to be able to express the orthogonal explicitly. For instance we invite the reader to think of the case where $\mathcal{H} = L^2[0, 1]$, $g(t) = e^t$, and $A = \{\alpha g : \ \alpha \in \mathbb{C}\}$; which functions belong to $A^\perp$? Other than saying that they are those functions $f$ in $L^2[0, 1]$ such that $\int_0^1 f(t)e^t\,dt = 0$, it does not look easy to come up with an explicit characterization.

**4.3.3. Definition.**

- *A subset $K \subset \mathcal{H}$ is said to be* ***convex*** *if for any $\xi, \eta \in K$, and for any $t \in [0,1]$, we have $t\xi + (1-t)\eta \in K$. This immediately implies (Exercise 4.3.1) that whenever $\xi_1, \ldots \xi_n \in K$ and $t_1, \ldots, t_n \in [0,1]$ with $t_1 + \cdots + t_n = 1$, we have $t_1\xi_1 + \cdots + t_n\xi_n \in K$. This latter expression is called a* ***convex combination*** *of elements in $K$.*
- *A subset $K \subset \mathcal{H}$ is a* ***subspace*** *if for any $\xi, \eta \in K$, and for any $s, t \in \mathbb{C}$, we have $s\xi + t\eta \in K$.*
- *The* ***span*** *of a subset $K \subset \mathcal{H}$ is the set*

$$\operatorname{span} K = \Big\{ \sum_{j=1}^{m} t_j \xi_j : \ m \in \mathbb{N}, t_1, \ldots, t_m \in \mathbb{C}, \ \xi_1, \ldots, \xi_m \in K \Big\},$$

*and its closure $\overline{\operatorname{span}} K$ is the* ***closed span*** *of $K$. So $\overline{\operatorname{span}} K$ is the* ***closed subspace generated by*** *$K$.*
- *The* ***distance*** *between $\xi \in \mathcal{H}$ and $K \subset \mathcal{H}$ is the number*

$$\operatorname{dist}(\xi, K) = \inf\{\|\xi - \eta\| : \ \eta \in K\}.$$

We do not require subspaces to be closed. Some authors save the word "subspace" for closed subspaces and use "linear manifold" for not necessarily closed subspaces, but there is no general agreement on this.

It is routine to check that

$$(\overline{\operatorname{span}} K)^{\perp} = K^{\perp}$$

(Exercise 4.3.11).

The next Lemma uses convexity and the parallelogram identity in a clever way. The manipulations in its proof are what distinguish Hilbert spaces among Banach spaces, and they play a crucial role in Propositions 4.3.6 and 4.3.8.

***4.3.4. Lemma.***

*Let $K \subset \mathcal{H}$ be a closed convex subset, and fix $\xi \in \mathcal{H}$. Then there exists $\nu_0 \in K$ such that $\operatorname{dist}(\xi, K) = \|\xi - \nu_0\|$, and $\nu_0$ is unique with that property.*

*When $K$ is a subspace, $\xi - \nu_0 \in K^{\perp}$.*

*Proof.* Let $\delta = \operatorname{dist}(\xi, K) = \inf\{\|\xi - \nu\| : \ \nu \in K\}$. Let $\{\nu_n\}_n \subset K$ be any sequence with $\|\xi - \nu_n\| \to \delta$. As $K$ is convex, $(\nu_n + \nu_m)/2 \in K$ and so

$$\|\xi - \nu_n + \xi - \nu_m\| = 2 \left\| \xi - \frac{\nu_n + \nu_m}{2} \right\| \geq 2\delta.$$

Now, using the Parallelogram Identity (4.2),

$$\begin{aligned}\|\nu_n - \nu_m\|^2 &= \|\xi - \nu_n - (\xi - \nu_m)\|^2 \\ &= 2\|\xi - \nu_n\|^2 + 2\|\xi - \nu_m\|^2 - \|\xi - \nu_n + \xi - \nu_m\|^2 \\ &\le 2\|\xi - \nu_n\|^2 + 2\|\xi - \nu_m\|^2 - 4\delta^2 \\ &\to 4\delta^2 - 4\delta^2 = 0\end{aligned}$$

as $n, m \to \infty$. So the sequence $\{\nu_n\}_n$ is Cauchy. Let $\nu_0 \in K$ be its limit; then $\|\xi - \nu_0\| = \delta$, as

$$\delta \le \|\xi - \nu_0\| \le \|\xi - \nu_n\| + \|\nu_n - \nu_0\| \xrightarrow[n\to\infty]{} \delta + 0 = \delta.$$

If $\nu_1 \in K$ is another point where the distance is achieved, we get from the inequalities above that

$$\|\nu_0 - \nu_1\|^2 \le 2\|\xi - \nu_0\|^2 + 2\|\xi - \nu_1\|^2 - 4\delta^2 = 4\delta^2 - 4\delta^2 = 0,$$

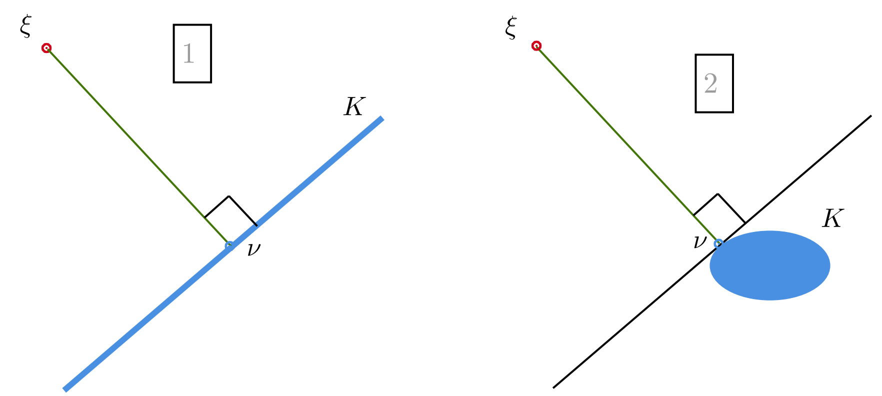
so $\nu_1 = \nu_0$.

When $K$ is a subspace and $\eta \in K$, we have $\|\xi - \nu_0\| \le \|\xi - \nu_0 + \eta\|$, since $\nu_0 - \eta \in K$. Squaring and expanding the right side of the inequality, we get

$$0 \le \|\eta\|^2 + 2\mathrm{Re}\,\langle \xi - \nu_0, \eta\rangle, \quad \eta \in K.$$

As this inequality holds for any $t\eta$ with $t \in \mathbb{C}$, it has to be that $\langle \xi - \nu_0, \eta\rangle = 0$ for all $\eta \in K$ (Exercise 4.3.4); that is, $\xi - \nu_0 \in K^\perp$. $\square$

*4.3.5. Example.* For some basic intuition on what Lemma 4.3.4 says, consider




In the situation $\boxed{1}$, the convex set $K$—in blue—is a subspace, while it is not a subspace in $\boxed{2}$. The black line in $\boxed{2}$ is the tangent space to $K$ at $\nu$. In both cases we can see how the green line $\xi - \eta$ realizes $\mathrm{dist}(\xi, K)$. The wonderful simplicity of the proof of Lemma 4.3.4 lets us know that the obvious idea in the figure above holds in any Hilbert space, even infinite-dimensional ones (where we have little to no geometrical intuition!). The ideas in Exercise 4.2.5 produce examples of normed spaces where the uniqueness in Lemma 4.3.4 fails. But the existence can fail, too (see Remark 5.1.15).

When our convex subset is a subspace, a lot more can be said, as we will see in the next two results.

**4.3.6. Proposition.**

*Let $A \subset \mathcal{H}$. Then $A^\perp$ is a closed subspace of $\mathcal{H}$, $A \cap A^\perp = \{0\}$ and $A^{\perp\perp} = \overline{\operatorname{span}} A$.*

*Proof.* Let $\xi, \eta \in A^\perp$, $a \in \mathbb{C}$. Then, for any $\nu \in A$ we have
$$\langle a\xi + \eta, \nu\rangle = a\langle \xi, \nu\rangle + \langle \eta, \nu\rangle = 0 + 0 = 0.$$
So $A^\perp$ is a vector space. If $\xi_n \in A^\perp$ for all $n$ and $\xi_n \to \xi$, then for any $\nu \in A$ we have—since the inner product is continuous (Exercise 4.3.8)—
$$\langle \xi, \nu\rangle = \lim_n \langle \xi_n, \nu\rangle = \lim_n 0 = 0.$$
So $A^\perp$ is closed.

The second assertion is trivial: if $\xi \in A \cap A^\perp$, then $\langle \xi, \xi\rangle = 0$, so $\xi = 0$.

For the third assertion, by definition $A \subset A^{\perp\perp}$, and so $\overline{\operatorname{span}} A \subset A^{\perp\perp}$ by the first part of the proof. Let $\xi \in A^{\perp\perp}$. As $\overline{\operatorname{span}} A$ is closed and convex, we get from Lemma 4.3.4 that the distance from $\xi$ to $\overline{\operatorname{span}} A$ is achieved at some $\nu_0 \in \overline{\operatorname{span}} A$, and $\xi - \nu_0 \in (\overline{\operatorname{span}} A)^\perp = A^\perp$. So $\xi - \nu_0 \in A^\perp \cap A^{\perp\perp} = \{0\}$. Then $\xi = \nu_0 \in \overline{\operatorname{span}} A$, and $A^{\perp\perp} = \overline{\operatorname{span}} A$. ☐

**4.3.7. Definition.**

*Let $\mathcal{K} \subset \mathcal{H}$ be a closed subspace and $\xi \in \mathcal{H}$. The unique $\nu_0 \in \mathcal{K}$ from Lemma 4.3.4 that realizes $\operatorname{dist}(\xi, \mathcal{K})$ is called the* ***projection*** *of $\xi$ on $\mathcal{K}$, denoted $P_\mathcal{K}\xi$.*

When $\mathcal{K}$ is a subspace, the assignment $\xi \to P_\mathcal{K}\xi$ is actually a linear map. Below we will actually construct $P_\mathcal{K}$ and show that it equals $\nu_0$.

**4.3.8. Proposition.**

*Let $\mathcal{K} \subset \mathcal{H}$ be a closed subspace. Then $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$ as a direct sum of subspaces. In particular, there exists a unique linear function $P_\mathcal{K} : \mathcal{H} \to \mathcal{H}$ such that*

*(i) $P_\mathcal{K}\xi \in \mathcal{K}$ for all $\xi \in \mathcal{H}$;*

*(ii)* $P_{\mathcal{K}}\nu = \nu$ *for all* $\nu \in \mathcal{K}$*;*

*(iii)* $\|P_{\mathcal{H}}\xi\| \leq \|\xi\|$*,* $\xi \in \mathcal{H}$*;*

*(iv)* $P_{\mathcal{K}^\perp}\xi = \xi - P_{\mathcal{K}}\xi$*,* $\xi \in \mathcal{H}$*;*

*(v)* $\ker P_{\mathcal{K}} = \mathcal{K}^\perp$*;*

*(vi)* $\operatorname{dist}(\xi, \mathcal{K}) = \|\xi - P_{\mathcal{K}}\xi\|$.

*Proof.* We first show that $\mathcal{K} + \mathcal{K}^\perp$ is closed. Since $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$, any $\xi \in \mathcal{K}+\mathcal{K}^\perp$ can be written in a unique way as $\xi_{\mathcal{K}}+\xi_\perp$ with $\xi_{\mathcal{K}} \in \mathcal{K}$, $\xi_\perp \in \mathcal{K}^\perp$; the uniqueness gives us that $(\xi + \eta)_{\mathcal{K}} = \xi_{\mathcal{K}} + \eta_{\mathcal{K}}$ and $(\xi + \eta)_\perp = \xi_\perp + \eta_\perp$. If $\{\xi_n\}$ is a Cauchy sequence in $\mathcal{K} + \mathcal{K}^\perp$, we write $\xi_n = \xi_{\mathcal{K},n} + \xi_{\perp,n}$ with $\xi_{\mathcal{K},n} \in \mathcal{K}$, $\xi_{\perp,n} \in \mathcal{K}^\perp$ for each $n$. As

$$\|\xi_{\mathcal{K},n} - \xi_{\mathcal{K},m}\|^2 \leq \|\xi_{\mathcal{K},n} - \xi_{\mathcal{K},m}\|^2 + \|\xi_{\perp,n} - \xi_{\perp,m}\|^2 = \|\xi_n - \xi_m\|^2,$$

we deduce that $\{\xi_{\mathcal{K},n}\}$ is Cauchy in $\mathcal{K}$, and thus convergent, and similarly for $\{\xi_{\perp,n}\}$; so $\{\xi_n\}$ converges to an element in $\mathcal{K} + \mathcal{K}^\perp$, and $\mathcal{K} + \mathcal{K}^\perp$ is closed.

We have $(\mathcal{K} + \mathcal{K}^\perp)^\perp = \mathcal{K}^\perp \cap \mathcal{K}^{\perp\perp} = \{0\}$ by Exercise 4.3.9 and Proposition 4.3.6. So $(\mathcal{K} + \mathcal{K}^\perp)^{\perp\perp} = \{0\}^\perp = \mathcal{H}$. As $\mathcal{K} + \mathcal{K}^\perp$ is a closed subspace, we get $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$ by Proposition 4.3.6.

Fix $\xi \in \mathcal{H}$. By the previous paragraph there exist unique $\xi_{\mathcal{K}} \in \mathcal{K}$, $\xi_\perp \in \mathcal{K}^\perp$ with $\xi = \xi_{\mathcal{K}}+\xi_\perp$. So we can define $P_{\mathcal{K}}\xi = \xi_{\mathcal{K}}$, and we are done with (i) and (ii). Linearity of $P_{\mathcal{K}}$ follows from the uniqueness in the decomposition $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$. The uniqueness of $P_{\mathcal{K}}$, in turn, also follows from the fact that it is the identity on $\mathcal{K}$ and zero on $\mathcal{K}^\perp$.

We have $\|\xi\|^2 = \|P_{\mathcal{K}}\xi + \xi_\perp\|^2 = \|P_{\mathcal{K}}\xi\|^2 + \|\xi_\perp\|^2$, so $\|P_{\mathcal{H}}\xi\| \leq \|\xi\|$. Since $\xi = P_{\mathcal{K}}\xi + \xi_\perp$, we have $P_{\mathcal{K}^\perp}\xi = \xi_\perp = \xi - P_{\mathcal{K}}\xi$.

We leave the proof of $\ker P_{\mathcal{K}} = \mathcal{K}^\perp$ as an exercise to the reader (Exercise 4.3.13).

Finally, for any $\eta \in \mathcal{K}$,

$$\|\xi - \eta\|^2 = \|\xi_{\mathcal{K}} - \eta + \xi_\perp\|^2 = \|\xi_{\mathcal{K}} - \eta\|^2 + \|\xi_\perp\|^2 \geq \|\xi_\perp\|^2.$$

That is, $\|\xi - \eta\| \geq \|\xi_\perp\|$ for all $\eta \in \mathcal{K}$, and the equality is achieved with $\eta = \xi_{\mathcal{K}}$. Thus

$$\operatorname{dist}(\xi, \mathcal{K}) = \|\xi_\perp\| = \|\xi - P_{\mathcal{K}}\xi\|. \qquad \square$$

*4.3.9. Remark*. To get an intuition on what Proposition 4.3.8 says, we suggest taking $\mathcal{H} = \mathbb{C}^2$, $\mathcal{K} = \{(z, 0) : z \in \mathbb{C}\}$, $\xi = (1, 1)$, and then calculating $\mathcal{K}^\perp$ and $P_{\mathcal{K}}$ explicitly. After doing this algebraically, one can look again at the picture $\boxed{1}$ in Example 4.3.5. Same with $\mathcal{K} = \{(z, z) : z \in \mathbb{C}\}$ and $\xi = (1, 0)$.

We now have the tools to focus on the representation of elements using a basis. In the results below, we consider infinite sums (i.e., series) of vectors in a Hilbert space $\mathcal{H}$. Two things need to be remembered about these. First, that an infinite series is not a sum, but a limit of sums; this implies the need for a topology. With number series this is a non-issue, because only one topology is usually considered on the real line or complex plane, and it comes from the natural order; but that is not the case in general. The other consideration is that we allow the set $J$ to be uncountable; the limit is then taken over the net of all finite subsets of $J$, ordered by inclusion. Namely,

**4.3.10. Definition.**

*Let $J$ be a set, $\eta \in \mathcal{H}$, and $\{\eta_j\}_{j\in J} \subset \mathcal{H}$. We write $\eta = \sum_{j\in J} \eta_j$ to mean that for each $\varepsilon > 0$ there exists a finite set $F_0 \in J$ with the property that for every finite set $F \in J$ with $F_0 \subset F$,*

$$\Big\|\eta - \sum_{j\in F} \eta_j\Big\| < \varepsilon.$$

*We call this **unconditional convergence**.*

Since the definition does not require any ordering in $J$, this notion of convergence does not agree with the usual calculus definition of series, even when $\mathcal{H} = \mathbb{C}$. When restricted to $\mathbb{C}$ and sequences, this definition would give us absolute convergence. One way to remark this difference is to notice, as the name indicates, that unconditional convergence is immune to reordering (as it should!):

***4.3.11. Lemma.***

*Let $J$ be a set and $\sigma : J \to J$ a bijection. Let $\mathcal{H}$ be a Hilbert space and $\{\eta_j\}_{j\in J} \subset \mathcal{H}$. Then $\sum_j \eta_j$ converges if and only if $\sum_j \eta_{\sigma(j)}$ converges, and in such case both series converge to the same element.*

*Proof.* Because $\sigma^{-1}$ is also a bijection, only one implication needs to be proven. Assume that $\sum_j \eta_j = \eta$. Fix $\varepsilon > 0$. Then there exists $F_0 \subset J$, finite, such that

$$\Big\|\eta - \sum_{j\in F} \eta_j\Big\| < \varepsilon.$$

for all finite $F$ with $J \supset F \supset F_0$. For any $G \supset \sigma^{-1}(F_0)$,

$$\Big\|\eta - \sum_{j\in G} \eta_{\sigma(j)}\Big\| = \Big\|\eta - \sum_{j\in \sigma(G)} \eta_j\Big\| < \varepsilon,$$

since $\sigma(G) \supset F_0$. Then $\sum_j \eta_{\sigma(j)} = \eta$. $\square$

In a complete space, a series is unconditionally convergent if and only if the tails of the series are small. Namely, $\sum_J \eta_j$ converges if and only if for each $\varepsilon > 0$ there exists finite $F_0 \subset J$ such that

$$\Big\| \sum_{j \in F \setminus F_0} \eta_j \Big\| < \varepsilon \tag{4.9}$$

for all finite $F \supset F_0$. Indeed, if $\sum_j \eta_j = \eta$, we take $F_0$ as in the definition of unconditional convergence for $\frac{\varepsilon}{2}$. Then for any finite $F \supset F_0$,

$$\Big\| \sum_{j \in F \setminus F_0} \eta_j \Big\| = \Big\| \eta - \sum_{j \in F_0} \eta_j - \Big( \eta - \sum_{j \in F} \eta_j \Big) \Big\| < 2\frac{\varepsilon}{2} = \varepsilon.$$

Conversely, if (4.9) holds for all $\varepsilon > 0$, we can get finite sets $F_0 \subset F_1 \subset F_2 \subset \cdots \subset J$ such that

$$\Big\| \sum_{j \in F_n \setminus F_0} \eta_j \Big\| < \frac{1}{n}.$$

This gives us that the sequence $\Big\{ \sum_{j \in F_n} \eta_j \Big\}_n$ is Cauchy, since for $m \geq n$

$$\Big\| \sum_{j \in F_m} \eta_j - \sum_{j \in F_n} \eta_j \Big\| = \Big\| \sum_{j \in F_m \setminus F_0} \eta_j - \sum_{j \in F_n \setminus F_0} \eta_j \Big\| < \frac{1}{m} + \frac{1}{n}.$$

So the sequence converges to some $\eta \in \mathcal{H}$ since the space is complete. Now (4.9) allows us to show that $\eta = \sum_{j \in J} \eta_j$.

*4.3.12. Remark*. Let us emphasize how unconditional convergence is different from the usual way series are defined in calculus. When we write

$$\eta = \lim_{m \to \infty} \sum_{j=1}^{m} \eta_j,$$

the existence of this limit does not necessarily imply unconditional convergence. Typical example of this is the dimension 1 case where $\sum_n \frac{(-1)^n}{n}$ converges in this latter sense but not unconditionally. This is what is usually called *conditional convergence*, and it has the pathology that it is not immune to reordering. In the finite-dimensional setting, unconditional convergence is equivalent to absolute convergence, but in general absolute converge is stronger than unconditional convergence. Unconditional convergence will be discussed in some detail in subsection 9.8.4. An infinite-dimensional Hilbert/Banach space always has an unconditionally convergent series that is not absolutely convergent. For a Hilbert space, this can be checked directly (see Example 4.3.19). For an arbitrary Banach space this is shown in [DR50b].

**4.3.13. Definition.**

*For a set $J$, we define*

$$\ell^2(J) = \Big\{ c : J \to \mathbb{C} : \sum_j |c(j)|^2 < \infty \Big\}. \tag{4.10}$$

We have already considered this space, in the case of the counting measure and $J = \mathbb{N}$, in Section 2.8. Where there is no ambiguity, we will use the more common notation $c_j$ instead of $c(j)$. But note that there **is** room for ambiguity in the next paragraph and in Lemma 4.3.20; the issue is how to denote a sequence of sequences, see the discussion before Lemma 4.3.20. We make $\ell^2(J)$ a complex vector space by using pointwise addition and multiplication by scalars. The first-time reader who is overwhelmed by the definitions above, should simply think that $J = \mathbb{N}$ for the time being. That is actually true for the vast majority of appearances of Hilbert spaces, both in this book and in the literature.

Even when $J$ is not countable in (4.10), the set $\{j : \ c(j) \neq 0\}$ is countable (Exercise 4.3.5).

We proved in Theorem 2.8.12 that $\ell^2(J)$ is complete as a normed space when given the norm $\|c\| = \Big(\sum_j |c(j)|^2\Big)^{1/2}$.

Another basic fact that we will use repeatedly is that the norm is continuous. Namely, if $\lim_j \eta_j = \eta$, then $\|\eta\| = \lim_j \|\eta_j\|$. This follows from the triangle inequality:

$$\big| \, \|\eta\| - \|\eta_j\| \, \big| \leq \|\eta - \eta_j\|.$$

**4.3.14. Definition.**

*An* ***orthonormal set*** *in $\mathcal{H}$ is a subset $\{\xi_j\}_{j\in J} \subset \mathcal{H}$ such that*

*(i) $\xi_k \perp \xi_j$ whenever $k \neq j$;*

*(ii) $\|\xi_j\| = 1$ for all $j \in J$.*

*When only (i) holds, we say that $\{\xi_j\}$ is* ***orthogonal****. An* ***orthonormal basis*** *for $\mathcal{H}$ is a maximal orthonormal set, where maximal means $\{\xi_j : j \in J\}^\perp = \{0\}$.*

An immediate consequence of orthonormality, that we will use very often and without mention, is the "Pythagorean Theorem":

***4.3.15. Lemma.***

*Let $\mathcal{H}$ be a Hilbert space, $\xi_1, \ldots, \xi_n \in \mathcal{H}$ orthonormal, and $c_1, \ldots, c_n \in \mathbb{C}$. Then*

$$\Big\| \sum_{j=1}^n c_j \xi_j \Big\|^2 = \sum_{j=1}^n |c_j|^2.$$

*Proof.* Exercise 4.3.2. □

***4.3.16. Theorem.***

*Every orthonormal set can be extended to an orthonormal basis. Thus, every Hilbert space has an orthonormal basis.*

*Proof.* Let $\{\xi_j\}_{j\in J}$ be an orthonormal set in $\mathcal{H}$. Let

$$\mathcal{F} = \big\{ \{\eta_k\}_{k\in K} : \text{ orthonormal, } \{\xi_j\}_{j\in J} \subset \{\eta_k\}_{k\in K} \big\}.$$

Consider in $\mathcal{F}$ the order given by inclusion. Suppose that $\{F_\ell\}_{\ell\in L}$ is a chain in $\mathcal{F}$. Put $F_0 = \bigcup_\ell F_\ell$; then $F_0$ is orthonormal and it contains $\{\xi_j\}_{j\in J}$, so $F_0 \in \mathcal{F}$. Also $F_\ell \subset F_0$ for all $\ell$, so $F_0$ is an upper bound for the chain. As we have shown that every chain in $\mathcal{F}$ admits an upper bound in $\mathcal{F}$, Zorn's Lemma guarantees that there exists a maximal element $F \in \mathcal{F}$. □

*4.3.17. Examples.*

The importance of orthonormal bases is given by the fact that they allow us to represent any vector in $\mathcal{H}$ as an "infinite linear combination" of elements in the basis. This is not obvious to mere mortals, and will require a proof (cfr. Theorem 4.3.21).

(a) In $\mathbb{C}^n$, the **canonical basis** is the set $\{e_1, \ldots, e_n\}$, where $e_k$ is the vector with a 1 in the $k^{\text{th}}$ entry and zeros elsewhere.

(b) More generally, in $\ell^2(\mathbb{N})$ the **canonical basis** is given by $\{\xi_n\}_{n\in\mathbb{N}}$, where $\xi_k(j) = \delta_{k,j}$.

(c) In $L^2(\mathbb{T})$ with Lebesgue measure (where $\mathbb{T}$ is the unit circle in the complex plane), an orthonormal basis is given by $\{z \longmapsto z^k\}_{k\in\mathbb{Z}}$ (this is the **Fourier Basis**).

(d) Orthonormal bases are very non-unique, even if there are some preferred ones like those mentioned above. Given any linearly independent set, the Gram–Schmidt process (Exercise 4.3.3) will produce an orthonormal set, which can then extended to a basis. Geometrically, if you rotate an orthonormal basis you will obtain an orthonormal basis; this gives us an intuitive proof that any Hilbert space of any dimension has uncountably many orthonormal bases. At the simplest, the one-dimensional Hilbert space $\mathbb{C}$ has bases $\{\lambda\}$ for each $\lambda \in \mathbb{T}$. In $\mathbb{C}^2$ we can still write all orthonormal bases explicitly: an orthonormal basis in $\mathbb{C}^2$ is necessarily of the form

$$\big\{(e^{ia}\cos t, e^{ib}\sin t), (e^{ic}\sin t, e^{id}\cos t)\big\}, \tag{4.11}$$

where $a, b, c \in \mathbb{R}$, $t \in [0, \pi/2]$ can be arbitrary and

$$d = \pi - a + b + c$$

The canonical basis is the case $t = a = b = d = 0$, $c = -\pi$.

Further down the road we will see that the proof of Lemma 4.3.18 shows that a Hilbert space of dimension $|J|$ is isometrically isomorphic to $\ell^2(J)$ (Theorem 4.4.4).

***4.3.18. Lemma.***

*Let $\{\xi_j\}_{j\in J}$ be an orthonormal set in $\mathcal{H}$, and $c = \{c_j\}_{j\in J} \subset \mathbb{C}$. Then $\sum_j c_j\xi_j$ converges if and only if $c \in \ell^2(J)$. Moreover, in this case,*

$$\|\sum_j c_j\xi_j\| = \|c\|_2.$$

*Proof.* Assume first that $\sum_j c_j\xi_j$ converges. Given $\varepsilon > 0$ there exists a finite set $F_0 \subset J$, such that $\left\|\sum_{j\in F\setminus F_0} c_j\xi_j\right\| < \varepsilon$ for all finite $F \supset F_0$. Using that $\{\xi_j\}$ is orthonormal,

$$\sum_{j\in F\setminus F_0} |c_j|^2 = \sum_{j\in F\setminus F_0} \|c_j\xi_j\|^2 = \Big\|\sum_{j\in F\setminus F_0} c_j\xi_j\Big\|^2 < \varepsilon^2.$$

We conclude that $\sum_j |c_j|^2 < \infty$. The same equality shows that if we now assume that $\sum_j |c_j|^2 < \infty$, then $\sum_j c_j\xi_j$ converges.

When $\sum_j c_j\xi_j$ converges, using that the norm is continuous we get

$$\Big\|\sum_j c_j\xi_j\Big\|^2 = \lim_F \Big\|\sum_{j\in F} c_j\xi_j\Big\|^2 = \lim_F \sum_{j\in F} |c_j|^2 = \sum_j |c_j|^2. \qquad \square$$

*4.3.19. Example.* We mentioned on Remark 4.3.12 that an infinite-dimensional Hilbert space always has unconditionally convergent series that are not absolutely convergent. The easiest example is, for any countable orthonormal set $\{\xi_n\}$,

$$\sum_{n=1}^{\infty} \frac{1}{n}\,\xi_n.$$

This series does not converge absolutely, as $\sum_n \left\|\frac{1}{n}\,\xi_n\right\| = \sum_n \frac{1}{n} = \infty$. But it converges unconditionally by Lemma 4.3.18.

As with the proof of completeness above, Lemma 4.3.20 below is a source of confusion for those new to functional analysis. In the particular case where $J = \mathbb{N}$, the Lemma deals with sequences of sequences. The confusion will hopefully be less with the emphasis on seeing the elements in $\ell^2(J)$ as functions instead of nets. With enough practice, reading, and re-reading, one eventually gets familiar with these kind of things.

***4.3.20. Lemma.***

*Let $\{\xi_j\}_{j\in J}$ be an orthonormal set in $\mathcal{H}$, and let $\{c_n\}_{n\in\mathbb{N}} \subset \ell^2(J)$ be a sequence. Then the following statements are equivalent:*

*(i)* $c_n \xrightarrow[n\to\infty]{} c$ *in* $\ell^2(J)$;

*(ii)* $\displaystyle\sum_{j\in J} c_n(j)\xi_j \xrightarrow[n\to\infty]{} \sum_{j\in J} c(j)\xi_j$ *in* $\mathcal{H}$.

*Proof.* We have

$$\begin{aligned}\Big\|\sum_{j\in J} c_n(j)\xi_j - \sum_{j\in J} c(j)\xi_j\Big\|^2 &= \Big\|\sum_{j\in J}(c_n(j)-c(j))\xi_j\Big\|^2 = \sum_{j\in J}|c_n(j)-c(j)|^2 \\ &= \|c_n - c\|_2^2.\end{aligned}$$

Note that we are using Lemma 4.3.18 in the second equality, and also to guarantee that the series are well-defined. □

We are finally in position to assert that given an orthonormal basis, any element can be represented in terms of it.

***4.3.21. Theorem.***

*Let $\mathcal{H}$ be a Hilbert space, $\{\xi_j\}_{j\in J}$ an orthonormal basis. Fix $\xi \in \mathcal{H}$. Then there exist unique coefficients $\{c_j\} \subset \mathbb{C}$ with $c \in \ell^2(J)$ and such that*

$$\xi = \sum_{j\in J} c_j\,\xi_j.$$

*Moreover, for all $j$ we have $c_j = \langle \xi, \xi_j \rangle$.*

*Proof.* Let $K \subset \mathcal{H}$ be the set

$$K = \Big\{ \sum_{j\in J} c_j\xi_j : \ c \in \ell^2(J) \Big\}.$$

This is well-defined by Lemma 4.3.18. We claim that $K$ is a closed subspace and that $K^\perp = \{0\}$. Then Proposition 4.3.6 shows that $K = \mathcal{H}$.

That $K$ is a subspace follows from the linearity of the limit. To see that $K^\perp$ is trivial, since $\{\xi_j : \ j \in J\} \subset K$, we have that $K^\perp \subset \{\xi_j : \ j \in J\}^\perp = \{0\}$.

The fact that $K$ is closed follows from Lemma 4.3.20—that shows that $K$ is homeomorphic to $\ell^2(J)$—and the fact that $\ell^2(J)$ is complete (Theorem 2.8.12).

The last equality follows from the continuity of the inner product and the orthogonality of the $\xi_j$. □

From Theorem 4.3.21 we now know that every element in a Hilbert space can be written as an "infinite linear combination" of elements from an orthonormal basis. The price to pay is that to sum infinite series we need to incorporate a topology. One could of course try to avoid that, and deal with infinite dimensional vector spaces as just that, vector spaces that have an infinite basis. Such a thing always exists—at least in the usual set theory framework of ZFC—and it is called a **Hamel basis**: a subset $\{\xi_j\} \subset \mathcal{H}$ such that every element in $\mathcal{H}$ is a linear combination of elements in $\{\xi_j\}$. Hamel bases are really not useful, in the sense that it is usually impossible to express them explicitly: any Hamel basis of any complete infinite-dimensional vector space is uncountable (this is a consequence of Baire's Category Theorem 6.3.1; see Exercise 6.3.3).

*4.3.22. Remark*. Suppose that $\mathcal{H}$ is a Hilbert space and $K \subset \mathcal{H}$ is a dense subspace. Does $\mathcal{H}$ have an orthonormal basis made out of elements of $K$? The answer is yes, when $\mathcal{H}$ is separable. If $\{\eta_n\} \subset K$ is dense in $\mathcal{H}$, we may replace $\{\eta_n\}$ with a linearly independent subset, and then perform Gram–Schmidt on this subset. This gives us an orthonormal set in $\mathcal{H}$ such that its linear span is dense, which means that we get an orthonormal basis. When $\mathcal{H}$ is not separable, though, the above can fail (Exercise 4.3.19).

We finish with two straightforward—and very useful!—consequences of Theorem 4.3.21: Bessel's Inequality, and Parseval's Equality.

***4.3.23. Corollary.***

*Let $\mathcal{H}$ be a Hilbert space, $\{\xi_j\}_j$ an orthonormal set. Then, for any $\xi \in \mathcal{H}$,*

$$\sum_j |\langle \xi, \xi_j\rangle|^2 \leq \|\xi\|^2 \qquad \textit{(Bessel's Inequality)}. \tag{4.12}$$

*If moreover $\{\xi_j\}_j$ is an orthonormal basis, then*

$$\sum_j |\langle \xi, \xi_j\rangle|^2 = \|\xi\|^2 \qquad \textit{(Parseval's Equality)} \tag{4.13}$$

*and*

$$\langle \xi, \eta\rangle = \sum_j \langle \xi, \xi_j\rangle\langle \xi_j, \eta\rangle, \qquad \eta \in \mathcal{H}. \tag{4.14}$$

*Proof.* If $\{\xi_j\}_j$ is already an orthonormal basis, then (4.13) follows directly from Lemma 4.3.18 and Theorem 4.3.21. When the orthonormal set is not necessarily a basis, we can extend it to a basis (Theorem 4.3.16) and get the equality in (4.13); the left-hand-side in (4.12) is then a series of non-negative terms with less terms that the one in (4.13), which proves the inequality.

The proof of (4.14) is left to the reader (Exercise 4.3.16). ☐

The equality (4.14) generalizes the usual $a \cdot b = \sum_k a_k \overline{b_k}$ in $\mathbb{C}^n$.

**4.3.1. Exercises.**

(4.3.1) *(Definition of convexity)* Let $K \subset \mathcal{H}$. Show that the following statements are equivalent:

(i) for any $\xi, \eta \in K$ and $t \in [0,1]$, $t\xi + (1-t)\eta \in K$;

(ii) for any $\xi_1, \ldots, \xi_n \in K$ and $t_1, \ldots, t_n \in [0,1]$ with $\sum_j t_j = 1$, we have $\sum_j t_j\xi_j \in K$.

(4.3.2) Prove Lemma 4.3.15.

(4.3.3) Let $\eta_1, \eta_2, \ldots, \eta_n$ be linearly independent. Show that if $\xi_1 = \eta_1/\|\eta_1\|$ and

$$\xi_{k+1} = \frac{\eta_{k+1} - \sum_{j=1}^k \langle \eta_{k+1}, \xi_j\rangle\, \xi_j}{\|\eta_{k+1} - \sum_{j=1}^k \langle \eta_{k+1}, \xi_j\rangle\, \xi_j\|} \tag{4.15}$$

then $\xi_1, \ldots, \xi_n$ is orthonormal and

$$\operatorname{span}\{\xi_1, \ldots, \xi_n\} = \operatorname{span}\{\eta_1, \ldots, \eta_n\}.$$

Show that the same algorithm works if the original set of vectors is countably infinite. This process of "orthonormalizing a basis" is called the **Gram–Schmidt Process**.

(4.3.4) Let $\mathcal{K} \subset \mathcal{H}$ be a subspace, and let $\xi \in \mathcal{H}$. Prove that if

$$0 \leq \|\nu\|^2 + 2\mathrm{Re}\,\langle \xi, \nu \rangle, \quad \nu \in \mathcal{K},$$

then $\xi \in \mathcal{K}^\perp$.

(4.3.5) Let $J$ be a set and $c : J \to \mathbb{C}$ a function. If $\sum_j |c(j)|^2 < \infty$, show that $\{j \in J : \ c(j) \neq 0\}$ is countable.

(4.3.6) Show that any orthonormal basis of $\mathbb{C}^2$ is of the form shown in (4.11).

(4.3.7) For each of the Hilbert spaces considered in Examples 4.3.17, find examples of orthonormal bases other than the canonical ones.

(4.3.8) Prove that the inner product in a Hilbert space is jointly continuous in its two variables.

(4.3.9) Let $\mathcal{K}_1, \mathcal{K}_2 \subset \mathcal{H}$ be subspaces. Show that

$$(\mathcal{K}_1 + \mathcal{K}_2)^\perp = \mathcal{K}_1^\perp \cap \mathcal{K}_2^\perp.$$

Show also that the equality is not true for arbitrary subsets.

(4.3.10) Let $A \subset \mathcal{H}$. Use Proposition 4.3.8 to obtain directly that $A^{\perp\perp} = \overline{\mathrm{span}}\,A$.

(4.3.11) Let $A \subset \mathcal{H}$. Show that $(\overline{\mathrm{span}}\,A)^\perp = A^\perp$.

(4.3.12) Let $\{\mathcal{H}_j\}$ be a family of closed subspaces of $\mathcal{H}$. Show that

$$\Big(\bigcup_j \mathcal{H}_j\Big)^\perp = \bigcap_j \mathcal{H}_j^\perp, \qquad\qquad \Big(\bigcap_j \mathcal{H}_j\Big)^\perp = \overline{\mathrm{span}} \bigcup_j \mathcal{H}_j^\perp.$$

(4.3.13) Show that if $\mathcal{K} \subset \mathcal{H}$ is a closed subspace, then $\ker P_\mathcal{K} = \mathcal{K}^\perp$.

(4.3.14) Let $A \subset \mathcal{H}$. Prove that $A^{\perp\perp}$ is equal to the intersection of all closed subspaces of $\mathcal{H}$ that contain $A$.

(4.3.15) Let $\mathcal{K} \subset \mathcal{H}$ be a subspace. Show that $\mathcal{K}$ is dense in $\mathcal{H}$ if and only if $\mathcal{K}^\perp = \{0\}$.

(4.3.16) Prove the equality (4.14).

(4.3.17) Let $A \subset \mathcal{H}$ be a set and $f : \mathcal{H} \to \mathcal{H}$ a function such that $f(\mathcal{H}) \subset A$ and $\xi - f(\xi) \in A^\perp$ for all $\xi \in \mathcal{H}$. Show that $A$ is a closed subspace and that $f = P_A$.

(4.3.18) Let $\{\xi_n\}_n \subset \mathcal{H}$ be an orthonormal basis, and put

$$M = \Big\{\eta \in \mathcal{H} : \ \sum_{n=1}^\infty \Big(1 + \frac{1}{n}\Big)^2 \left|\langle \eta, \xi_n \rangle\right|^2 \leq 1\Big\}.$$

Show that $M$ is bounded, closed, convex, and that it has no element with greatest norm. *(This gives us an example of a bounded, closed, convex subset and an $\mathbb{R}$-valued nonnegative continuous function that does not attain its maximum. Of course, $M$ is not compact).*

(4.3.19) Show an example of a non-separable Hilbert space $\mathcal{H}$ with a dense subspace $K$ such that $K$ contains no orthonormal basis. *(The examples known to the author are not trivial)*

(4.3.20) Let $\mathcal{H}$ be a Hilbert space. Since $\mathbb{R} \subset \mathbb{C}$, we can see $\mathcal{H}$ as a real vector space, that we denote $\mathcal{H}_\mathbb{R}$.

(i) Show that $\mathcal{H}_\mathbb{R}$ is a real Hilbert space with the inner product $\langle \xi, \eta \rangle = \operatorname{Re} \langle \xi, \eta \rangle$.

(ii) Show that if $\{\xi_j\}$ is an orthonormal basis for $\mathcal{H}$, then $\{\xi_j\} \cup \{i\xi_j\}$ is an orthonormal basis for $\mathcal{H}_\mathbb{R}$.

(iii) We say that $\xi, \eta \in \mathcal{H}$ are **real orthogonal**, denoted $\xi \perp_\mathbb{R} \eta$, if $\operatorname{Re} \langle \xi, \eta \rangle = 0$. Show that $\xi \perp \eta$ if and only if $\xi \perp_\mathbb{R} \eta$ and $i\xi \perp_\mathbb{R} \eta$.

(iv) Show that if $H \subset \mathcal{H}$ is a real subspace, then $(iH)^\perp_\mathbb{R} = iH^\perp_\mathbb{R}$.

## 4.4. Dimension

In a vector space, the dimension is defined as the cardinality of a(ny) basis. For an infinite-dimensional Hilbert space a basis in the usual algebraic sense (a maximal linearly independent set of vectors) is always uncountable, and there is no possibility of writing such a basis explicitly. Because of that, and because we have orthonormal bases available, it makes a lot more sense to consider dimension for Hilbert spaces in terms of cardinalities of orthonormal bases. The first step is to check that such a definition makes sense:

***4.4.1. Proposition.***

*Let $\mathcal{H}$ be a Hilbert space. Then any two orthonormal bases of $\mathcal{H}$ have the same cardinality.*

*Proof.* If $\mathcal{H}$ has a finite orthonormal basis, the result is well-known from linear algebra. So let $\{\xi_j\}_{j\in J}$, $\{\eta_\ell\}_{\ell\in L}$ be two orthonormal bases for $\mathcal{H}$, with $L$ and $J$ infinite. By Theorem 4.3.21 there exist, for each $j \in J$, coefficients $\{c_{j,\ell}\}_{\ell\in L}$ such that $\xi_j = \sum_{\ell\in L} c_{j,\ell}\eta_\ell$. As $1 = \|\xi_j\| = \sum_{\ell\in L} |c_{j,\ell}|^2$, each set $L_j = \{\ell \in L : c_{j,\ell} \neq 0\}$ is countable (Exercise 4.3.5). The closed subspace

$$\overline{\operatorname{span}}\Big\{\eta_\ell : \ \ell \in \bigcup_{j\in J} L_j\Big\}$$

contains all $\xi_j$, and thus it equals $\mathcal{H}$. So $\{\eta_\ell\}_{\ell\in\cup_{j\in J}L_j}$ is an orthonormal basis for $\mathcal{H}$, and thus $L = \bigcup_{j\in J} L_j$. As each $L_j$ is countable, we get $|L| \leq |J|$ by Corollary 1.6.34. Switching roles we get $|J| \leq |L|$, and so $|J| = |L|$ by Schröeder–Bernstein (Theorem 1.6.13). $\square$

In view of Proposition 4.4.1,

**4.4.2. Definition.**

*Let $\mathcal{H}$ be a Hilbert space. The* ***dimension*** *of $\mathcal{H}$, denoted $\dim \mathcal{H}$, is the cardinality of any orthonormal basis of $\mathcal{H}$.*

Dimension characterizes Hilbert spaces completely; for this, we need to talk about isomorphisms. As usual, a homomorphism is a function that preserves the structure; for Hilbert spaces this is not really standard terminology, no one talks about "homomorphisms of Hilbert spaces", but we make the definition below to be able emphasize how rigid the structure of a Hilbert space is. In the case of a Hilbert space, the "structure" consists of the two vector space operations and the inner product.

**4.4.3. Definition.**

*Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. A* ***homomorphism of Hilbert spaces*** *is a linear map $V : \mathcal{H}_1 \to \mathcal{H}_2$ that preserves the inner product, i.e.*

$$\langle V\xi, V\eta\rangle_2 = \langle \xi, \eta\rangle_1.$$

It follows immediately that $V$ is an isometry, that is $\|V\xi\| = \|\xi\|$, and so every homomorphism is injective; conversely, the polarization identity guarantees that any linear isometry is a homomorphism. Thus the **isomorphisms of Hilbert spaces** are precisely the surjective linear isometries. Such a map is usually called a **unitary** map. We will meet them again in Chapter 10, after a couple mentions in Chapters 8 and 9. A particular consequence of this is that isomorphisms of Hilbert spaces coincide with *isometric* isomorphisms of Hilbert spaces.

***4.4.4. Theorem.***

*Let $\mathcal{H}$ be a Hilbert space of dimension $|J|$ for some set $J$. Then $\mathcal{H} \simeq \ell^2(J)$.*

*Proof.* Let $\{\xi_j\}_{j\in J}$ be an orthonormal basis for $\mathcal{H}$. Define $V : \mathcal{H} \to \ell^2(J)$ by

$$V\xi = \{\langle \xi, \xi_j\rangle\}_j.$$

This is linear. If we take $c(j) = \langle \xi, \xi_j\rangle$, Parseval's Equality (4.13) guarantees that $c \in \ell^2(J)$ and that $\|V(\xi)\|_2 = \|c\|_2 = \|\xi\|$. So $V$ does map into $\ell^2(J)$, and it is an isometry. Then polarization makes $V$ a homomorphism. Finally, given $c \in \ell^2(J)$ we have $\xi = \sum_j c_j\xi_j \in \mathcal{H}$, and $V(\xi) = c$; so $V$ is onto. $\square$

***4.4.5. Corollary.***

*Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Then $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic if and only if* $\dim \mathcal{H}_1 = \dim \mathcal{H}_2$.

*Proof.* A homomorphism preserves orthonormality; being surjective, it also preserves orthogonals, and so it maps orthonormal bases to orthonormal bases; thus it preserves dimension. Conversely, if $\dim \mathcal{H}_1 = \dim \mathcal{H}_2 = |J|$, then by Theorem 4.4.4 we have

$$\mathcal{H}_1 \simeq \ell^2(J) \simeq \mathcal{H}_2.$$

□

So dimension is a complete invariant for a Hilbert space; for each cardinality $\alpha$, there is a unique Hilbert space $\mathcal{H}$ with $\dim \mathcal{H} = \alpha$. In the case of countable dimension, we can relate dimension with a topological property. Recall that a topological space is **separable** if it contains a dense countable subset. It is straightforward to check that finite-dimensional Hilbert spaces are separable (Exercise 4.4.2).

***4.4.6. Proposition.***

*Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Then the following statements are equivalent:*

*(i) $\mathcal{H}$ has countable infinite dimension;*

*(ii) $\mathcal{H}$ is separable.*

*Proof.* Exercise 4.4.3. □

The next result is true in more generality than Hilbert spaces, but we can get a very simple and direct proof in our current setting.

**4.4.7. Definition.**

*The* ***closed unit ball*** *of $\mathcal{H}$ is the set $(\mathcal{H})_1 = \{\xi \in \mathcal{H} : \|\xi\| \leq 1\}$. The (open)* ***ball of radius*** *$r$ around $\xi \in \mathcal{H}$ is the set*

$$B_r(\xi) = \{\eta \in \mathcal{H} : \|\xi - \eta\| < r\}.$$

*So* $(\mathcal{H})_1 = \overline{B_1(0)}$.

Having proven Proposition 4.4.6, this next proof is natural.

*4.4.8. Theorem.*

*Let $\mathcal{H}$ be a Hilbert space. Then $\mathcal{H}$ is finite-dimensional if and only if $(\mathcal{H})_1$ is compact.*

*Proof.* If $\mathcal{H}$ is finite-dimensional, we can use the compacity of the closed unit disk in $\mathbb{C}$ to deduce the compacity of the unit ball in $\mathcal{H}$ (Exercise 4.4.4).

Conversely, if $\mathcal{H}$ is infinite-dimensional, fix an orthonormal basis $\{\xi_j\}$. For any two distinct indices $k, j$,

$$\|\xi_k - \xi_j\|^2 = \|\xi_k\|^2 + \|\xi_j\|^2 - 2\mathrm{Re}\,\langle \xi_k, \xi_j\rangle = \|\xi_k\|^2 + \|\xi_j\|^2 = 2.$$

This implies that the sequence $\{\xi_j\}$ admits no convergent subsequence, and thus $(\mathcal{H})_1$ is not compact. □

The reader who is encountering Theorem 4.4.8 for the first time would do well in pondering a little about this. In the space where one usually learns analysis, $\mathbb{R}^n$, "compact" is synonym with "closed and bounded"—this is the Heine–Borel theorem. But here, closed balls in an infinite-dimensional Hilbert space are closed and bounded, but not compact. Compactness is a tricky matter in infinite-dimension. It is as much if not more useful than in finite dimension; compactness is one of the reasons why we will sometimes work with other topologies where closed balls are compact. We will discuss such topologies in Chapter 7 and beyond (most details will appear in sections Sections 7.2 and 12.1, but these topologies are used for many results).

### 4.4.1. Exercises.

(4.4.1) Write a complete proof of Corollary 4.4.5, without using Theorem 4.4.4. *(Hint: show that the linear operator induced by mapping one orthonormal basis to another is an isomorphism)*

(4.4.2) Show that a finite-dimensional Hilbert space is separable.

(4.4.3) Prove Proposition 4.4.6.

(4.4.4) Show that the unit ball of a finite-dimensional Hilbert space is compact.

## 4.5. The Riesz Representation Theorem

While at this stage it is not clear why, we will now focus for a moment on linear functionals on a Hilbert space. A **linear functional** on $\mathcal{H}$ is a function $\varphi : \mathcal{H} \to \mathbb{C}$ that is linear (i.e. preserves addition and multiplication by scalars).

Linear functionals play an important role in Functional Analysis. We will see, however, that they are rather trivial in the case of Hilbert spaces. Their usefulness, for now, will be that they are a very particular kind of *linear operators*, and we can use them to start gaining knowledge about the issues with continuity that arise in infinite dimension. Eventually, linear functionals will play a central role for the whole rest of the book.

In finite dimension, it is easy to see that linearity implies continuity. But that is not the case in infinite dimension. For instance, let $\mathcal{H}_0 = \mathbb{C}[x]$ (polynomials), with $\langle p, q\rangle = \int_0^1 p(t)\overline{q(t)}\,dt$. Then $\mathcal{H}_0$ is not a Hilbert space (it is not complete), but its completion $L^2[0,1]$ is. On $\mathcal{H}_0$ we may define the linear functional $\varphi(p) = p'(1)$. This is linear, as both taking derivative and evaluating at a point are. Now, if $p_n(t) = t^n$, then

$$\|p_n\|_2 = \Big(\int_0^1 t^{2n}\,dt\Big)^{1/2} = \frac{1}{\sqrt{2n+1}}.$$

Thus $p_n \to 0$ in $\mathcal{H}$. Meanwhile, $\varphi(p_n) = n$ and hence $\varphi$ is not continuous at 0, and it cannot naturally be extended to the completion of $\mathcal{H}_0$ (the definition doesn't even make sense for an arbitrary continuous function, let alone a measurable one, and without continuity there is no obvious property to decide what the values at $\mathcal{H} \setminus \mathcal{H}_0$ would be, and why they would be meaningful). It probably feels a bit weird that the functional we constructed is only defined on a pre-Hilbert space, or on a dense subspace of a Hilbert space, but that turns out to be a fairly common feature of discontinuous functionals; that is, they are **densely defined**. It is possible to construct discontinuous linear functionals defined on all of $\mathcal{H}$—for any infinite-dimensional Hilbert space—but it is known that they cannot be expressed in terms of simple formulas as above.

To construct an example of a discontinuous linear functional defined in all of an infinite-dimensional Hilbert space $\mathcal{H}$, fix a linearly independent set $\{\xi_n\} \subset \mathcal{H}$ (an orthonormal basis, for example, but all we need is the linear independence). Define $\varphi(\xi_n) = n$. The set $\{\xi_n\}$ can be algebraically extended—via Zorn's Lemma—to a Hamel basis of $\mathcal{H}$, and we extend $\varphi$ by linearity as 0. Then $|\varphi(\xi_n/n)| = 1$ while $\|\xi_n/n\| \longrightarrow 0$, so $\varphi$ cannot be continuous.

So, how can we tell if a linear functional is continuous? Here is a characterization.

***4.5.1. Proposition.***

*Let $\varphi : \mathcal{H} \to \mathbb{C}$ be linear. Then the following statements are equivalent:*

*(i) $\varphi$ is continuous;*

*(ii) $\varphi$ is continuous at $0$;*

*(iii) there exists $\xi_0 \in \mathcal{H}$ such that $\varphi$ is continuous at $\xi_0$;*

*(iv) there exists $r > 0$ such that, for every $\xi \in \mathcal{H}$, $|\varphi(\xi)| \leq r\,\|\xi\|$.*

*Proof.* (i) $\Longrightarrow$ (ii): trivial.

(ii) $\Longrightarrow$ (iii): trivial.

(iii) $\Longrightarrow$ (iv): assume that $\varphi$ is continuous at $\xi_0$. Then there exists $\delta > 0$ such that, whenever $\|\xi - \xi_0\| < \delta$, then $|\varphi(\xi) - \varphi(\xi_0)| \leq 1$. If $\|\eta\| < \delta$, we get

$$|\varphi(\eta)| = |\varphi(\eta + \xi_0 - \xi_0)| = |\varphi(\eta + \xi_0) - \varphi(\xi_0)| \leq 1.$$

For any $\xi \in \mathcal{H}$, taking $\eta = \delta\xi/(2\|\xi\|)$, we have $\|\eta\| < \delta$ and then $|\varphi(\eta)| \leq 1$; that is,

$$\left|\varphi\left(\frac{\delta\xi}{2\|\xi\|}\right)\right| \leq 1,$$

from where we get $|\varphi(\xi)| \leq \frac{2}{\delta}\|\xi\|$.

(iv) $\Longrightarrow$ (i): If $\xi_n \to \xi$, this means that $\|\xi_n - \xi\| \to 0$. So

$$|\varphi(\xi_n) - \varphi(\xi)| = |\varphi(\xi_n - \xi)| \leq r\,\|\xi_n - \xi\| \to 0. \qquad \square$$

It is instructive to note that the proof of Proposition 4.5.1 does not really use that $\mathcal{H}$ is a Hilbert space; it only uses that it is a normed vector space. So the argument applies to normed spaces. And it also does not use that the codomain is $\mathbb{C}$, and any normed space could be considered in the codomain.

***4.5.2. Corollary.***

*Let $\mathcal{H}$ be a finite-dimensional normed space, $\varphi : \mathcal{H} \to \mathbb{C}$ linear. Then $\varphi$ is continuous.*

*Proof.* Let $\xi_1, \ldots, \xi_n \in \mathcal{H}$ be a basis of $\mathcal{H}$. Let $r = \max\{\|\xi_1\|, \ldots, \|\xi_n\|\}$. For any $\xi \in \mathcal{H}$, we have $\xi = \sum_{j=1}^n c_j\xi_j$. Then, using Cauchy–Schwarz,

$$\begin{aligned}|\varphi(\xi)| &= \left|\varphi\Big(\sum_{j=1}^n c_j\xi_j\Big)\right| = \left|\sum_{j=1}^n c_j\,\varphi(\xi_j)\right| \leq r\sum_{j=1}^n |c_j| \\ &\leq r\sqrt{n}\left(\sum_{j=1}^n |c_j|^2\right)^{1/2} = r\sqrt{n}\,\|\xi\|.\end{aligned}$$

By Proposition 4.5.1, the functional $\varphi$ is continuous. $\square$

Because of Proposition 4.5.1,

**4.5.3. Definition.**

*A continuous linear functional is called a* ***bounded linear functional****. More generally, we say that continuous linear map $T : \mathcal{X} \to \mathcal{Y}$ between normed spaces is* ***bounded****.*

It turns out that bounded linear functionals on a Hilbert space can be completely characterized (which is not the case with bounded functionals on more general spaces, as we will see later). Note first that if $\varphi : \mathcal{H} \to \mathbb{C}$ is given by $\varphi(\xi) = \langle \xi, \eta \rangle$ for some $\eta$, then

$$|\varphi(\xi)| = |\langle \xi, \eta \rangle| \leq \|\eta\| \, \|\xi\|, \qquad \xi \in \mathcal{H},$$

so $\varphi$ is bounded. The remarkable thing is that this is the only way to produce bounded linear functionals on a Hilbert space:

***4.5.4. Theorem. (Riesz Representation Theorem)***

*Let $\varphi : \mathcal{H} \to \mathbb{C}$, linear and bounded. Then there exists a unique $\eta \in \mathcal{H}$ such that $\varphi(\xi) = \langle \xi, \eta \rangle$ for all $\xi \in \mathcal{H}$.*

*Proof.* The uniqueness is direct: if both $\eta, \eta' \in \mathcal{H}$ satisfy the statement, then $\langle \xi, \eta - \eta' \rangle = 0$ for all $\xi \in \mathcal{H}$; in particular, $\|\eta - \eta'\| = 0$. So we move to consider the existence. If $\varphi = 0$ identically, we take $\eta = 0$. So we can assume that $\varphi \neq 0$; since $\ker \varphi = \varphi^{-1}(\{0\})$ is a proper closed subspace of $\mathcal{H}$ (this is where we use that $\varphi$ is bounded!), there exists $\eta_1 \in (\ker \varphi)^\perp$. Necessarily, $\varphi(\eta_1) \neq 0$, so by replacing $\eta_1$ with a multiple if needed, we may assume that $\varphi(\eta_1) = 1$. Let $\xi \in \mathcal{H}$; since $\xi - \varphi(\xi)\eta_1 \in \ker \varphi$, we have $\langle \xi - \varphi(\xi)\eta_1, \eta_1 \rangle = 0$. Then, if $\eta = \eta_1 / \|\eta_1\|^2$,

$$\varphi(\xi) = \left\langle \xi, \frac{\eta_1}{\langle \eta_1, \eta_1 \rangle} \right\rangle = \langle \xi, \eta \rangle. \qquad \square$$

The study of the space of bounded linear functionals (that is, the **dual**) of a space is a staple of Functional Analysis, as we will see from Section 5.5 onwards. For now we will use linear functionals for the following characterization of Hilbert spaces.

***4.5.5. Proposition.***

*Let $\mathcal{H}_0$ be a pre-Hilbert space. If $M = M^{\perp\perp}$ for all closed subspaces $M \subset \mathcal{H}_0$, then $\mathcal{H}_0$ is a Hilbert space.*

*Proof.* All orthogonals are taken with respect to $\mathcal{H}_0$. We want to show that $\mathcal{H}_0$ is complete. Let $\mathcal{H}$ be the completion of $\mathcal{H}_0$. Fix $\nu \in \mathcal{H}$, nonzero, and let

$$M = \{\xi \in \mathcal{H}_0 : \ \langle \xi, \nu \rangle = 0\}.$$

We have $M = \mathcal{H}_0 \cap \ker \varphi$, where $\varphi(\xi) = \langle \xi, \nu \rangle$. As $\varphi$ is bounded, $\ker \varphi$ is closed in $\mathcal{H}$, and $M$ is closed in $\mathcal{H}_0$. By hypothesis, $M^{\perp\perp} = M$. It follows that $M^{\perp} \neq \{0\}$; because otherwise $M = M^{\perp\perp} = \{0\}^{\perp} = \mathcal{H}_0$, and this contradicts $\nu \neq 0$ by the density of $\mathcal{H}_0$ in $\mathcal{H}$. Let $\eta \in M^{\perp} \setminus \{0\}$, and let $\psi(\xi) = \langle \xi, \eta \rangle$. As $\eta$ is nonzero and $\eta \in \mathcal{H}_0 \setminus M$, we have $\varphi(\eta) \neq 0$. Multiplying by a suitable scalar, we may assume that $\varphi(\eta) = 1$. For any $\xi \in \mathcal{H}_0$, the element $\xi - \varphi(\xi)\eta$ is in $\mathcal{H}_0 \cap \ker \varphi = M$. Let $c = \psi(\eta) = \|\eta\|^2 > 0$. As $\eta \in M^{\perp}$, we have

$$0 = \langle \xi - \varphi(\xi)\eta, \eta \rangle = \psi(\xi - \varphi(\xi)\eta) = \psi(\xi) - c\varphi(\xi) = \langle \xi, \eta - c\nu \rangle, \qquad \xi \in \mathcal{H}_0.$$

But $\mathcal{H}_0$ is dense in $\mathcal{H}$, so $\langle \xi, \eta - c\nu \rangle = 0$ for all $\xi \in \mathcal{H}$. Taking in particular $\xi = \eta - c\nu$, we get that $\eta - c\nu = 0$. As $c \neq 0$, this implies that $\nu \in \mathcal{H}_0$, and thus $\mathcal{H}_0 = \mathcal{H}$. □

The key idea in the proof of Proposition 4.5.5 is that if $\ker \varphi \subset \ker \psi$, then $\psi$ is a multiple of $\varphi$. Later, in Lemma 5.5.10, we will prove a more general version of this fact, that will be used repeatedly.

### 4.5.1. Exercises.

(4.5.1) Let $\varphi$ be a bounded functional on $\mathcal{H}$. Fix an orthonormal basis $\{\xi_j\}$. Given any $\xi = \sum_j c_j \xi_j \in \mathcal{H}$, show that $\varphi(\xi) = \sum_j c_j \varphi(\xi_j)$ and use this fact to get an alternative proof of the Riesz Representation Theorem (4.5.4). Explain where the continuity of $\varphi$ was used.

(4.5.2) The vector space of all bounded functionals on $\mathcal{H}$ is called its **dual**, denoted by $\mathcal{H}^*$. Show that

$$\|\varphi\| = \sup\{|\varphi(\xi)| : \ \|\xi\| = 1\}$$

defines a norm on $\mathcal{H}^*$, and that if $\eta \in \mathcal{H}$ is the vector corresponding to $\varphi$ via the Riesz Representation Theorem, then $\|\varphi\| = \|\eta\|$.

(4.5.3) Let $\varphi : \mathcal{H} \to \mathbb{C}$ be nonzero, linear, and bounded. Show that

$$\dim(\ker \varphi)^{\perp} = 1.$$

(4.5.4) Let $\varphi : \mathcal{H} \to \mathbb{C}$ be linear. Prove that the following two statements are equivalent:

(i) $\varphi$ is bounded;

(ii) $\ker \varphi$ is closed.

Use the ideas in your proof to show that if $\varphi$ is unbounded, then $\ker \varphi$ is dense in $\mathcal{H}$.

CHAPTER

# 5

# Banach spaces

From an abstract point of view, Hilbert spaces are normed vector spaces that are complete, and such that the norm comes from an inner product. We will now remove that last condition, and what remains is what we call a Banach space. Banach spaces appear naturally in many contexts, and we will encounter them in many forms throughout the rest of the text.

As with Hilbert spaces, one can consider both real and complex Banach spaces. We will mostly consider complex spaces, though most results and examples work both in the complex and in the real case.

## 5.1. Normed spaces

We have already met norms in Chapter 4, and now they are going to be the main tool at our disposal. So here is a formal definition.

**5.1.1. Definition.**

*A **norm** on a complex vector space $\mathcal{X}$ is a function $p : \mathcal{X} \to [0, \infty)$ such that*

*(i) For all $x \in \mathcal{X}$, $p(x) = 0$ implies $x = 0$.*

*(ii) For all $\lambda \in \mathbb{C}$, for all $x \in \mathcal{X}$, $p(\lambda x) = |\lambda|\, p(x)$. (**homogeneity**)*

*(iii)* *For $x, y \in \mathcal{X}$, $p(x+y) \leq p(x)+p(y)$.* (***subadditivity/Triangle Inequality***)

*A* ***seminorm*** *is a function $p : \mathcal{X} \to [0, \infty)$ satisfying (ii) and (iii).*

*Two (semi)norms $p, q$ on $\mathcal{X}$ are said to be* ***equivalent*** *if there exist $\alpha, \beta \in \mathbb{R}_+$ with*

$$\alpha\, q(x) \leq p(x) \leq \beta\, q(x), \ \forall x \in \mathcal{X}.$$

In most cases a norm is denoted by $\|x\| = p(x)$; sometimes with subscripts to indicate a specific kind of norm. Equivalence of norms is an equivalence relation (Exercise 5.1.3). Recall the notation $B_r^{\mathcal{X}}(x)$ to indicate the open ball of radius $r$ around $x$ in $\mathcal{X}$.

A useful application of the triangle inequality is the **reverse triangle inequality** (Exercise 1.8.24): for any seminorm $p$,

$$|p(x) - p(y)| \leq p(x - y).$$

From now on we will consider vector spaces $\mathcal{X}$ that are equipped with a norm. The norm induces a metric via $d(x, y) = \|x - y\|$, and so $\mathcal{X}$ can also be seen as a topological space. The properties of the norm guarantee that the native operations from $\mathcal{X}$, namely addition of vectors and multiplication of a scalar by a vector, are both continuous (Exercise 5.1.1).

**5.1.2. Definition.**

*A* ***Banach space*** *is a complete normed space.*

Any Banach space has a Hamel basis—assuming the Axiom of Choice, which we do—but this is useless in practice. Hilbert spaces have a more useful kind of "basis" where we can represent elements as series of the form coefficient-times-basis element. Not every Banach space has something like that, but many do. In such cases, we talk about a **Schauder basis**. Properly,

**5.1.3. Definition.**

*Let $\mathcal{X}$ be a Banach space. A* ***Schauder basis*** *for $\mathcal{X}$ is a countable set $\{e_k\} \subset \mathcal{X}$ such that for every $x \in \mathcal{X}$ we have unique coefficients $\{c_k\}$ with $x = \sum_{k=1}^{\infty} c_k e_k$.*

Here are some of the most common examples of Banach spaces. At this point we offer no proof that each is indeed a Banach space.

*5.1.4. Example*. The one-dimensional complex vector space $\mathbb{C}$ is a Banach space with the norm $|a+ib| = (a^2+b^2)^{1/2}$.

*5.1.5. Example.* Any finite-dimensional normed space is a Banach space, and any basis is a Schauder basis. For instance $\mathbb{C}^n$ with the Euclidean norm is a Banach space (this is actually $L^2(\{1,\dots,n\})$ with the counting measure, in language from Section 2.8. We can make $\mathbb{C}^n$ into a different Banach space by using another norm; for instance $\|x\|_1 = \sum_j |x_j|$, or $\|x\|_\infty = \max\{|x_1|,\dots,|x_n|\}$.

*5.1.6. Example.* Any Hilbert space is a Banach space, and any orthonormal basis is a Schauder basis.

*5.1.7. Example.* For any $p \in \mathbb{R}$ with $1 \le p \le \infty$,

$$\ell^p(\mathbb{N}) = \{\alpha : \mathbb{N} \to \mathbb{C} : \ \|\alpha\|_p < \infty\}$$

is a Banach space with the norm

$$\|\alpha\|_p = \begin{cases} \Big(\displaystyle\sum_{k=1}^{\infty} |\alpha(k)|^p\Big)^{1/p} & \text{if } p < \infty \\ \sup\{|\alpha(k)| : \ k \in \mathbb{N}\} & \text{if } p = \infty \end{cases}$$

These spaces carry a natural Schauder basis when $p < \infty$, the **canonical basis** $\{e_n\}$ where $e_n(k) = \delta_{n,k}$. We have $\alpha = \sum_n \alpha(n)\, e_n$, with the series converging in the $p$-norm. The space $\ell^\infty(\mathbb{N})$, on the other hand, does not have a Schauder basis: we can write elements formally as "series" over the canonical basis, but these series do not converge in the infinity norm. This is not just an irrelevant distinction; for instance we have that in $\ell^\infty(\mathbb{N})$, formally, $1 = \sum_n e_n$, but

$$\Big\|1 - \sum_{n=1}^{m} e_n\Big\| = 1 \qquad \text{for all } m.$$

*5.1.8. Example.* More generally, as considered in Section 2.8, given any measure space $(X, \Omega, \mu)$, the space

$$L^p(X,\Omega,\mu) = \{f : X \to \mathbb{C} : \ f \text{ is measurable and } \int_{\mathcal{X}} |f|^p \, d\mu < \infty\}$$

where $1 \le p < \infty$, is a Banach space with the norm

$$\|f\|_p = \left(\int_X |f|^p \, d\mu\right)^{1/p}$$

It is not obvious whether $L^p$ spaces have a Schauder basis in general; they do for some measure spaces.

For $p = \infty$, we define $L^\infty(X, \Omega, \mu)$ to be the set of essentially bounded functions, with the norm

$$\|f\|_\infty = \operatorname*{ess\,sup}_X |f|.$$

*5.1.9. Example.* Let $c_0 = \{\alpha : \mathbb{N} \to \mathbb{C} : \lim_{k\to\infty} \alpha(k) = 0\}$ with $\|\cdot\|_\infty$. Then $c_0$ is a Banach space; in fact, it is a closed subspace of $\ell^\infty(\mathbb{N})$. The canonical basis is a Schauder basis for $c_0$ (Exercise 5.1.9).

*5.1.10. Example.* Let $T$ be a compact Hausdorff topological space. Then $C(T)$, the space of continuous functions on $T$, is a Banach space with the norm $\|f\|_\infty = \sup\{|f(t)| : t \in T\}$. For the particular case where $T = [0, 1]$, an example of a Schauder basis can be seen in Examples 9.8.4.

*5.1.11. Example.* Let $T$ be a locally compact Hausdorff topological space. Then $C_0(T)$, the space of continuous functions on $T$ that vanish at infinity, is a Banach space with the norm $\|\cdot\|_\infty$. The space $C_b(T)$, of continuous bounded functions on $T$, is also a Banach space with the infinity norm.

*5.1.12. Example.* Note that the inclusion of Banach spaces $c_0 \subset \ell^\infty(\mathbb{N})$ is a particular case (when $T = \mathbb{N}$ with the discrete topology) of $C_0(T) \subset C_b(T)$.

*5.1.13. Example.* We can also get finite-dimensional examples of Banach spaces by taking $\mathcal{X} = \mathbb{C}^n$, and using one of the norms $\|\cdot\|_p$, $1 \le p \le \infty$.

*5.1.14. Example. (another finite-dimensional example)* Let $\mathcal{X} = \{p \in \mathbb{C}[x] : \deg p \le 7\}$, with $\|p\| = \max\{|p(t)| : \in [0, 1]\}$.

This list of examples should not be considered exhaustive by any means. The zoo of Banach spaces is extremely large. We will develop a few more examples in Section 5.6.

*5.1.15. Remark.* We saw in Lemma 4.3.4 that in a Hilbert space distances to closed subspaces are attained; that is not the case for Banach spaces in general. Consider the Banach space $L^1[0, 1]$, and let

$$K = \Big\{g \in L^1[0, 1] : \int_0^1 tg(t)\, dt = 0\Big\}.$$

Then $K$ is a closed subspace. Indeed, if $g_n \to g$ with $g_n \in K$ for all $n$, then

$$\begin{aligned}\Big|\int_0^1 t\,g(t)\,dt\Big| &\le \Big|\int_0^1 t\,\big(g(t)-g_n(t)\big)\,dt\Big| + \Big|\int_0^1 t\,g_n(t)\,dt\Big| \\ &= \Big|\int_0^1 t\big(g(t)-g_n(t)\big)\,dt\Big| \le \|g-g_n\|_1.\end{aligned}$$

As this can be done for any $n$, we get that the integral for $g$ is zero, and so $g \in K$ and $K$ is closed. Now we consider $\operatorname{dist}(f,K)$, where $f$ is the constant function $f(t)=1$. For any $g \in K$,

$$\begin{aligned}\|g-f\|_1 = \int_0^1 |g-f| &\ge \int_0^1 t\,|(g(t)-f(t)|\,dt \\ &\ge \Big|\int_0^1 t\,(g(t)-f(t))\,dt\Big| = \int_0^1 t\,dt = \frac{1}{2}.\end{aligned}$$

If we take $g_n = 1 - \frac{n^2}{2n-1}\,1_{[1-1/n,1]}$, then

$$\begin{aligned}\int_0^1 t\,g_n(t)\,dt &= \frac{1}{2} - \int_0^1 t\,\frac{n^2}{2n-1}\,1_{[1-1/n,1]}(t)\,dt \\ &= \frac{1}{2} - \frac{n^2}{2n-1}\int_{1-1/n}^1 t\,dt = \frac{1}{2} - \frac{1}{2} = 0,\end{aligned}$$

so $g_n \in K$ and

$$\|g_n - f\|_1 = \int_0^1 \frac{n^2}{2n-1}\,1_{[1-1/n,1]}(t)\,dt = \frac{n}{2n-1} \xrightarrow[n\to\infty]{} \frac{1}{2}.$$

This establishes that $\operatorname{dist}(f,K) = \frac{1}{2}$. And this distance is not attained. Indeed, suppose that $\|f-g\|_1 = \frac{1}{2}$ for some $g \in K$. This gives us equality in the inequalities above, so in particular

$$\int_0^1 (1-t)\,|g-f| = 0,$$

which forces $g=f$ a.e., since the integrand is non-negative. But then $g=1$, which is not in $K$, a contradiction.

We finish the section with an elementary but deep result that will be of use later. One can safely skip to Section 5.2.

**5.1.16. Definition.**

*Let $\mathcal{X}$ be a normed space, $z \in \mathcal{X}$. The* ***reflection*** *of $\mathcal{X}$ on $z$ is the map $S_z : \mathcal{X} \to \mathcal{X}$ given by $S_z x = 2z - x$.*

*A (not necessarily linear) map $T : \mathcal{X} \to \mathcal{Y}$ is an* ***isometry*** *if $\|Tx - Ty\| = \|x-y\|$ for all $x, y \in \mathcal{X}$.*

*A map $T : \mathcal{X} \to \mathcal{Y}$ is **affine** if*

$$T(tx + (1-t)z) = tTx + (1-t)Tz, \qquad x, z \in \mathcal{X},\ 0 \le t \le 1.$$

Direct calculations (Exercise 5.1.10) show that $S_z^2 = \mathrm{id}_{\mathcal{X}}$, that $S_z$ is an isometry, that $S_z$ affine, that $z$ is the only fixed point of $S_z$, and that

$$\|z - S_z x\| = \|z - x\|, \qquad \|x - S_z x\| = 2\|x - z\|, \quad x, z \in \mathcal{X}. \tag{5.1}$$

Most isometries that we will consider will be linear, but there is a small exception in the very near future. We will see that they are always almost linear, though. The following characterization will be used without mention:

**5.1.17. Lemma.**

*Let $\mathcal{X}, \mathcal{Y}$ be vector spaces and $T : \mathcal{X} \to \mathcal{Y}$ a function. The following statements are equivalent:*

*(i) there exists $L : \mathcal{X} \to \mathcal{Y}$, linear, and $y_0 \in \mathcal{Y}$ with $T = y_0 + L$;*

*(ii) $T - T0$ is real linear;*

*(iii) $T$ is affine.*

*Proof.* (i) $\Longrightarrow$ (ii) Note that $T0 = y_0$. For $t \in \mathbb{R}$ and $x, y \in \mathcal{X}$,

$$(T - y_0)(tx + y) = L(tx + y) = tLx + Ly = t(T - y_0)x + (T - y_0)y,$$

so $T - T_0$ is real linear.

(ii) $\Longrightarrow$ (iii) For $x, y \in \mathcal{X}$ and $t \in [0,1]$,

$$\begin{aligned} T(tx + (1-t)y) &= T0 + (T - T0)(t + x + (1-t)y) \\ &= T0 + t(T - T0)x + (1-t)(T - T0)y \\ &= T0 + tTx + (1-t)Ty - tT0 - (1-t)T0 \\ &= tTx + (1-t)Ty \end{aligned}$$

(iii) $\Longrightarrow$ (i) Let $y_0 = T0$, and $L = T - y_0$. Consider first $t \in [0,1]$, $x \in \mathcal{X}$. We have

$$L(tx) = -y_0 + T(tx + (1-t)0) = -y_0 + tTx + (1-t)T0 = tLx.$$

For $t > 1$, from above

$$t\,Lx = t\,L\Big(\frac{1}{t}\,(tx)\Big) = t\,\frac{1}{t}\,L(tx) = L(tx).$$

So $L(tx) = tLx$ for all $t \ge 0$ and $x \in \mathcal{X}$. Now for $x, y \in \mathcal{X}$, using that $L$ is affine,

$$L(x + y) = L\Big(\frac{1}{2}\,2x + \frac{1}{2}\,2y\Big) = \frac{1}{2}\,L(2x) + \frac{1}{2}\,L(2y) = \frac{1}{2}\,2Lx + \frac{1}{2}\,2Ly = Lx + Ly$$

and $L$ is additive. Finally, $L(-x) + Lx = L(-x + x) = L0 = 0$; thus $L(-x) = -Lx$ and now

$$L(-t)x) = L(t(-x)) = tL(-x) = -tLx,$$

showing that $L$ also spits negative coefficients. Thus $L$ is real linear. $\square$

The proof of Proposition 5.1.18 that we show below is due to Bogdan Nica [Nic12].

***5.1.18. Proposition. (Mazur–Ulam)***

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{X} \to \mathcal{Y}$ a (not necessarily linear) bijective isometry. Then $T$ is affine.*

*Proof.* We want to show that $T(tx + (1-t)y) = tTx + (1-t)Ty$ for all $x, y \in \mathcal{X}$ and $t \in [0,1]$. Initially we will assume that $t = 1/2$. Fix $x, y \in \mathcal{X}$ and let $z = \frac{1}{2}x + \frac{1}{2}y$. Let

$$d(T) = \Big\| T\Big(\frac{1}{2}x + \frac{1}{2}y\Big) - \frac{1}{2}Tx - \frac{1}{2}Ty \Big\|.$$

Then

$$\begin{aligned} d(T) &\leq \frac{1}{2}\Big\| T\Big(\frac{1}{2}x + \frac{1}{2}y\Big) - Tx \Big\| + \frac{1}{2}\Big\| T\Big(\frac{1}{2}x + \frac{1}{2}y\Big) - Ty \Big\| \\ &= \frac{1}{4}\|Ty - Tx\| + \frac{1}{4}\|Tx - Ty\| = \frac{1}{2}\|x - y\|. \end{aligned}$$

Let $S : \mathcal{X} \to \mathcal{Y}$ be given by $Sw = Tx + Ty - w$. Then $S$ is a bijective isometry with $Sx = Ty$, $Sy = Tx$. Put $R = T^{-1}ST$. Then $R$ is a bijective isometry and $Rw = T^{-1}(Tx + Ty - Tw)$. In particular $Rx = y$, $Ry = x$. We have

$$\begin{aligned} d(R) &= \Big\| R\Big(\frac{1}{2}x + \frac{1}{2}y\Big) - \frac{1}{2}Rx - \frac{1}{2}Ry \Big\| \\ &= \Big\| T^{-1}\Big(Tx + Ty - T\Big(\frac{1}{2}x + \frac{1}{2}y\Big)\Big) - \frac{1}{2}x - \frac{1}{2}y \Big\| \\ &= \Big\| Tx + Ty - T\Big(\frac{1}{2}x + \frac{1}{2}y\Big) - T\Big(\frac{1}{2}x - \frac{1}{2}y\Big) \Big\| \\ &= 2\Big\| T\Big(\frac{1}{2}x - \frac{1}{2}y\Big) - \frac{1}{2}Tx - \frac{1}{2}Ty \Big\| \\ &= 2d(T). \end{aligned}$$

If $d(T) > 0$, by iterating the construction we can get a surjective isometry $R$ with $d(R) > \frac{1}{2}\|x - y\|$, contradicting the initial estimate. It follows that $d(T) = 0$; that is,

$$T\Big(\frac{1}{2}x + \frac{1}{2}y\Big) = \frac{1}{2}Tx + \frac{1}{2}Ty.$$

Now we need to get the above for $t \neq \frac{1}{2}$. We will proceed by induction. Suppose for the inductive step that

$$T\Big(\frac{\ell}{2^k}\,x + \frac{2^k-\ell}{2^k}\,y\Big) = \frac{\ell}{2^k}\,Tx + \frac{2^k-\ell}{2^k}\,Ty, \qquad \ell = 0,\dots,2^k.$$

Then, assuming without loss of generality that $\ell \leq 2^k$ (otherwise we replace it with $2^{k+1}-\ell$)

$$\begin{aligned} T\Big(\frac{\ell}{2^{k+1}}\,x + \frac{2^{k+1}-\ell}{2^{k+1}}\,y\Big) &= T\Big[\frac{1}{2}\,\Big(\frac{\ell}{2^k}\,x + \frac{2^k-\ell}{2^k}\,y\Big) + \frac{1}{2}\,y\Big] \\ &= \frac{1}{2}\,T\Big(\frac{\ell}{2^k}\,x + \frac{2^k-\ell}{2^k}\,y\Big) + \frac{1}{2}\,Ty \\ &= \frac{1}{2}\,\Big(\frac{\ell}{2^k}\,Tx + \frac{2^k-\ell}{2^k}\,Ty\Big) + \frac{1}{2}\,Ty \\ &= \frac{\ell}{2^{k+1}}\,Tx + \frac{2^{k+1}-\ell}{2^{k+1}}\,Ty. \end{aligned}$$

So $T$ preserves convex combinations with dyadic convex coefficients. As $T$ is continuous (being an isometry), we get $T(tx+(1-t)y) = tTx + (1-t)Ty$ for all $t \in [0,1]$. □

**5.1.1. Exercises.**

(5.1.1) Prove that in a normed vector space, addition of vectors and multiplication by a scalar are continuous.

(5.1.2) Let $\mathcal{X}$ be a normed space, $x \in \mathcal{X}$, and $\{x_n\}$ a sequence such that $x_n \to x$. Show that $\|x_n\| \to \|x\|$.

(5.1.3) Show that equivalence of norms is an equivalence relation.

(5.1.4) Let $\mathcal{X}$ be a normed space and $\{x_n\} \subset \mathcal{X}$ a Cauchy sequence. Show that the sequence is bounded, that is there exists $c > 0$ such that $\|x_n\| \leq c$ for all $n$.

(5.1.5) Consider the real Banach spaces $X_1 = (\mathbb{R}^2, \|\cdot\|_1)$, $X_2 = (\mathbb{R}^2, \|\cdot\|_2)$, $X_3 = (\mathbb{R}^2, \|\cdot\|_\infty)$. Find and describe graphically the unit ball of each space (the use of real spaces is only to allow the possibility of drawing the unit balls).

(5.1.6) Show that the canonical basis is a Schauder basis for $\ell^p(\mathbb{N})$ when $1 \leq p < \infty$.

(5.1.7) Show that the canonical basis is not a Schauder basis for $\ell^\infty(\mathbb{N})$.

(5.1.8) Show that $\ell^\infty(\mathbb{N})$ is not separable and hence it has no Schauder basis.

(5.1.9) Show that the canonical basis is a Schauder basis for $c_0$.

(5.1.10) Let $\mathcal{X}$ be a normed space and $x, z \in \mathcal{X}$, with $S_z$ the reflection of $\mathcal{X}$ on $z$. That is, $S_z x = -(x-z) + z = 2z - x$. Show that

(i) $S_z^2 = \mathrm{id}_{\mathcal{X}}$;

(ii) $S_z$ is an isometry;

(iii) $S_z$ is affine;

(iv) $z$ is the only fixed point of $S_z$;

(v) We have the equalities

$$\|z - S_z x\| = \|z - x\|, \qquad \|x - S_z x\| = 2\|x - z\|, \quad x, z \in \mathcal{X}.$$

## 5.2. Finite-dimensional Banach spaces

As seen in the examples above, a vector space can possibly be equipped with different norms. Exercise 5.1.5 shows that these different norms may generate the same topology: it is clear that for any point inside an open disk in $\mathbb{R}^2$ one can put an open square around the point and still remain inside the disk; and vice versa. This can be expressed algebraically by the inequalities

$$\|x\|_\infty \leq \|x\|_1 \leq \sqrt{2}\,\|x\|_2 \leq 2\|x\|_\infty, \qquad x \in \mathbb{R}^2$$

that say that the three norms are equivalent to each other. A natural question is whether it is possible that the same topology is induced by non-equivalent norms. And, also, what necessary and sufficient conditions are there for that situation to occur (or not)? It turns out that this can be answered in a fairly straightforward way:

***5.2.1. Proposition.***

*Let $\mathcal{X}$ be a vector space with two norms $p_1$ and $p_2$. Then the following statements are equivalent:*

*(i) $p_1$ and $p_2$ are equivalent.*

*(ii) $p_1$ and $p_2$ induce the same topology;*

*Proof.* (i) $\implies$ (ii) By assumption, there exist $\alpha, \beta > 0$ with $\alpha p_1(x) \leq p_2(x) \leq \beta p_1(x)$, $x \in \mathcal{X}$. Write $B_r^k(x) = \{y \in \mathcal{X} : \ p_k(y - x) < r\}$ for the balls in each of the corresponding topologies. We need to show that any ball given by $p_2$ is open in the topology given by $p_1$, and viceversa; since the roles are reversible, only one of the inclusions is needed. Namely, we need show that if $x \in \mathcal{X}$ and $r > 0$, then for any $y \in B_r^2(x)$ there exists $s > 0$ such that $B_s^1(y) \subset B_r^2(x)$. Given $y \in B_r^2(x)$, since $p_2(y - x) < r$, there exists $r_0$ with $p_2(y - x) < r_0 < r$. Let $s = (r_0 - p_2(y - x))/\beta$. If $z \in B_s^1(y)$, then

$$p_2(z - x) \leq p_2(z - y) + p_2(y - x) \leq \beta p_1(z - y) + p_2(y - x)$$

$$\leq \beta s + p_2(y-x) = r_0 < r.$$

So $z \in B_r^2(x)$, and thus $B_s^1(y) \subset B_r^2(x)$. This shows that the $p_1$ topology is finer than the $p_2$ topology; by exchanging roles, we get that the two topologies are equal.

(ii) $\Longrightarrow$ (i) If the two norms induce the same topology, then $B_1^2(0)$ is open in the topology induced by $p_1$; therefore there exists $r > 0$ such that $B_r^1(0) \subset B_1^2(0)$. For any nonzero $x \in \mathcal{X}$, let $y = \frac{r}{2p_1(x)}\,x$. Then $p_1(y) < r$, and so $p_2(y) < 1$, which we may write as $p_2(x) \leq \frac{2}{r}\,p_1(x)$. By exchanging roles we get the other inequality. □

In the case of a finite-dimensional space, there is only one possible topology:

***5.2.2. Theorem.***

*Let $\mathcal{X}$ be a finite-dimensional vector space. Then any two norms on $\mathcal{X}$ are equivalent.*

*Proof.* Let $p$ be a norm on $\mathcal{X}$. We will show that it is equivalent with a well-chosen norm; this will be enough, since we already know that equivalence of norms is an equivalence relation. Fix a basis $\{f_1, \ldots, f_n\}$ of $\mathcal{X}$. Given $x \in \mathcal{X}$, it can be written in a unique way as $x = \sum_j x_j f_j$. Define

$$\|x\|_1 = \sum_{j=1}^n |x_j|. \tag{5.2}$$

Straightforward computations shows that this is a norm. Also,

$$p(x) = p\Big(\sum_{j=1}^n x_j f_j\Big) \leq \sum_{j=1}^n |x_j|\,p(f_j) \leq \beta\|x\|_1, \tag{5.3}$$

where $\beta = \max_j p(f_j)$. This in particular implies that the function $\eta : \mathbb{C}^n \to [0,\infty)$ given by $\eta : (c_1, \ldots, c_n) \mapsto p(\sum_{j=1}^n c_j f_j)$ is continuous: namely,

$$\begin{aligned}|\eta(c) - \eta(d)| &= \Big|p\Big(\sum_{j=1}^n c_j f_j\Big) - p\Big(\sum_{j=1}^n d_j f_j\Big)\Big| \leq p\Big(\sum_{j=1}^n (c_j - d_j) f_j\Big)\\ &\leq \sum_{j=1}^n |c_j - d_j|\,p(f_j) \leq \beta \sum_{j=1}^n |c_j - d_j| \leq \sqrt{n}\,\beta\,\|c - d\|.\end{aligned}$$

Let $B = \{c = (c_1, \ldots, c_n) \in \mathbb{C}^n : \sum_j |c_j| = 1\}$. This set is closed and bounded in the usual topology, and so it is compact by the Heine–Borel Theorem. Then the continuous function $\eta$ attains its maximum and minimum in $B$. In particular, there exists $c \in B$ such that

$$\eta(c) \leq \eta(v) \qquad \text{for all } v \in B.$$

Note that $0 \notin B$ and that $\eta(v) = 0$ implies that $v = 0$, so $\eta(c) > 0$. Now, for any nonzero $x = \sum_{j=1}^n x_j f_j \in \mathcal{X}$ we have $(x_1/\|x\|_1, \dots, x_n/\|x\|_1) \in B$ and so

$$\eta(c) \leq \eta\left(\frac{(x_1, \dots, x_n)}{\|x\|_1}\right) = \frac{1}{\|x\|_1} p(\sum_{j=1}^n x_j f_j) = \frac{1}{\|x\|_1} p(x).$$

Thus

$$\|x\|_1 \leq \frac{1}{\eta(c)} p(x). \tag{5.4}$$

Now the two inequalities (5.3) and (5.4) give the equivalence between $p$ and $\|\cdot\|_1$. □

***5.2.3. Corollary.***

*Any two norms on a finite-dimensional vector space induce the same topology.*

*Proof.* This is a direct application of Theorem 5.2.2 and Proposition 5.2.1. □

***5.2.4. Corollary.***

*If $\mathcal{X}$ is a finite-dimensional normed vector space, then its closed unit ball $\overline{B_r^{\mathcal{X}}(0)}$ is compact.*

*Proof.* Exercise 5.2.5. □

Recall that, for a linear transformation, we use the word *bounded* to mean continuous. The motivation for this was given in Proposition 4.5.1 and will be completely justified later in Proposition 6.1.2.

***5.2.5. Corollary.***

*Let $\mathcal{X}$, $\mathcal{Y}$ be normed spaces, with $\mathcal{X}$ finite-dimensional. Then $\mathcal{X}$ is complete, any subspace of $\mathcal{X}$ is closed, and any linear transformation $\mathcal{X} \to \mathcal{Y}$ is bounded.*

*Proof.* Since the result is about topological properties and we know that all norms in a finite-dimensional space induce the same topology (Corollary 5.2.3), we may use any norm we want. So fix a basis $f_1, \dots, f_k$ of $\mathcal{X}$ and, as in (5.2), let

$$\|x\|_1 = \sum_{j=1}^k |x(j)| \tag{5.5}$$

when $x = \sum_{j=1}^k x(j)\, f_j$. Any subspace of $\mathcal{X}$ is itself a finite-dimensional normed space; so it is enough to show that a finite-dimensional normed space $\mathcal{X}$ is complete. Since $\mathcal{X}$ is finite-dimensional, its closed unit ball is compact; in particular it is closed. Given a Cauchy sequence $\{x_n\} \subset \mathcal{X}$ there exists $r > 0$ with $\|x_n\|_1 < r$ for all $n$ (Exercise 5.1.4). For each $j = 1, \ldots, k$,

$$|x_n(j) - x_m(j)| \leq \|x_n - x_m\|_1,$$

and hence the sequence $\{x_n(j)\}_n$ is Cauchy in $\mathbb{C}$. So the sequence $\{x_n\}$ converges entrywise; that is, there exists $x \in \mathbb{C}^k$ with $\lim_n x_n(j) = x(j)$. Namely, we write $x = \sum_{j=1}^k x(j) f_j \in \mathcal{X}$. As

$$\|x_n - x\|_1 = \sum_{j=1}^{k} |x_n(j) - x(j)| \xrightarrow[n\to\infty]{} 0,$$

$x$ is the limit in $\mathcal{X}$ of the sequence $\{x_n\}$, and $\mathcal{X}$ is complete.

Given $T : \mathcal{X} \to \mathcal{Y}$ linear, let $\{f_1, \cdots, f_k\}$ be a basis for $\mathcal{X}$. For $v \in \mathcal{X}$, write $v = \sum_j v_j f_j$, and then

$$\|Tv\| = \|\sum_j v_j T f_j\| \leq \sum_j |v_j|\, \|Tf_j\| \leq \beta' \|v\|_1,$$

where $\beta' = \max \|Tf_j\|$. By Theorem 5.2.2, there exists a constant $\beta$ such that $\|\cdot\|_1 \leq \beta \|\cdot\|$. So $\|Tv\| \leq \beta\beta' \|v\|$, and $T$ is bounded. □

**5.2.6. Corollary.**

*Let $S$ be a finite set. The spaces $\ell^p(S)$, $p \in [1, \infty]$ are all mutually isomorphic. But for $p < \infty$, $\ell^p(S)$ is not isometrically isomorphic to $\ell^\infty(S)$.*

*Proof.* The mutual isomorphism follows directly from the fact that they are all $n$-dimensional complex vector spaces, where $n = |S|$. The isomorphisms and their inverses are continuous by Corollary 5.2.5.

As for isometric isomorphism,suppose that $\gamma : \ell^\infty(S) \to \ell^p(S)$ is an isometric isomorphism. In $\ell^\infty(S)$ we take $a = e_1$, $b = e_1 + e_2$; then $\|a + b\|_\infty = 2 = \|a\|_\infty + \|b\|_\infty$. After applying $\gamma$, since it is isometric, we get $x = \gamma(a), y = \gamma(b) \in \ell^p(S)$ with $\|x+y\|_p = \|x\|_p + \|y\|_p$. By Corollary 2.8.10, there exist $\alpha, \beta \in \mathbb{C}$ with $\alpha x = \beta y$. As $\gamma$ is an isomorphism this gives $\alpha a = \beta b$, which is not true. Therefore $\ell^\infty(S)$ is not isometrically isomorphic to $\ell^p(S)$ for any $p \in [1, \infty)$. □

It turns out that the spaces $\ell^p(S)$ and $\ell^{p'}(S)$ are not isometrically isomorphic for $p \neq p'$ finite. In fact, if there exists an isometric embedding $\ell^p(S_1) \to \ell^{p'}(S_2)$

then $p = 2$, $p'$ is even, and there is a relation between $|S_1|$ and $|S_2|$. We refer to [LV93] for details and references.

Same as with Hilbert spaces (Theorem 4.4.8), the converse of Corollary 5.2.4 also holds. In order to prove it, in an arbitrary Banach space we cannot repeat the argument we used in Theorem 4.4.8 to show that the unit ball cannot be compact in infinite dimension. The argument worked for Hilbert spaces because we had an orthonormal basis, which gave us a sequence of equidistant vectors. To achieve the same in a Banach space, Riesz's Lemma is the key. Before the lemma, we restate for normed spaces a definition we made in the context of Hilbert spaces:

**5.2.7. Definition.**

*Let $\mathcal{X}$ be a normed space. The **distance** between $x \in \mathcal{X}$ and $\mathcal{Y} \subset \mathcal{X}$ is the number*

$$\operatorname{dist}(x, \mathcal{Y}) = \inf\{\|x - y\| : \ y \in \mathcal{Y}\}.$$

***5.2.8. Lemma. (F. Riesz)***

*Let $\mathcal{X}$ be a normed vector space, $\mathcal{Y} \subset \mathcal{X}$ a closed proper subspace. Then, for any $\varepsilon > 0$, there exists $x \in \mathcal{X}$, $\|x\| = 1$ such that*

$$\operatorname{dist}(x, \mathcal{Y}) > 1 - \varepsilon.$$

*Proof.* Let $x_0 \in \mathcal{X} \setminus \mathcal{Y}$, and put $c = \operatorname{dist}(x_0, \mathcal{Y})$. As $\mathcal{Y}$ is closed, $c > 0$. Choose $\delta > 0$ such that $c/(c+\delta) > 1-\varepsilon$. By definition of distance, there exists $y_0 \in \mathcal{Y}$ such that $c \le \|x_0 - y_0\| < c + \delta$. Now put

$$x = \frac{x_0 - y_0}{\|x_0 - y_0\|}$$

Then, for any $y \in \mathcal{Y}$, $y_0 + \|x_0 - y_0\| y \in \mathcal{Y}$ and so $\|x_0 - (y_0 + \|x_0 - y_0\| y)\| \ge c$; thus

$$\begin{aligned} \|x - y\| &= \left\| \frac{x_0 - y_0 - \|x_0 - y_0\| y}{\|x_0 - y_0\|} \right\| = \frac{\|x_0 - (y_0 + \|x_0 - y_0\| y)\|}{\|x_0 - y_0\|} \\ &> \frac{c}{c + \delta} > 1 - \varepsilon. \end{aligned}$$

□

***5.2.9. Theorem.***

*Let $\mathcal{X}$ be a normed vector space. Then $\mathcal{X}$ is finite-dimensional if and only if its unit ball is compact.*

*Proof.* If $\mathcal{X}$ is finite-dimensional, its unit ball is compact by Corollary 5.2.4.

Conversely, if $\mathcal{X}$ is infinite dimensional, start with $x_1 \in \mathcal{X}$ such that $\|x_1\| = 1$. Let $\mathcal{X}_1 = \mathbb{C}x_1$. We will construct an increasing sequence of finite-dimensional subspaces $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \cdots$ where all the inclusions are proper. We start with $\mathcal{X}_1$ and assume for induction that we have $\mathcal{X}_n = \operatorname{span}\{x_1, \ldots, x_n\}$, where $\|x_j\| = 1$, and $\|x_k - x_j\| > 1/2$ if $k \neq j$. By Corollary 5.2.5, the subspace $\mathcal{X}_n$ is closed. And, since $\mathcal{X}$ is infinite-dimensional, the inclusion $\mathcal{X}_n \subset \mathcal{X}$ is proper. So, by Lemma 5.2.8 there exists $x_{n+1} \in \mathcal{X}$ such that $\|x_{n+1}\| = 1$ and $\operatorname{dist}(x_{n+1}, \mathcal{X}_n) > 1/2$. We let $\mathcal{X}_{n+1} = \operatorname{span}\{\mathcal{X}_n \cup \{x_{n+1}\}\}$ and the induction is complete. The sequence $\{x_n\} \subset \mathcal{X}$ satisfies $\|x_n - x_m\| > 1/2$ for all $n \neq m$. Thus the unit ball of $\mathcal{X}$ cannot be compact. $\square$

**5.2.1. Exercises.**

(5.2.1) Show that a finite-dimensional normed space $\mathcal{X}$ is separable.

(5.2.2) By Theorem 5.2.2, the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are equivalent on $\mathbb{C}^n$. Find specific constants that realize the inequalities.

(5.2.3) Show that $\|\cdot\|_1$, defined in (5.2), is a norm.

(5.2.4) Let $\mathcal{X}$ be a normed vector space of dimension $n$. Prove that there is a bicontinuous isomorphism between $\mathcal{X}$ and $\mathbb{C}^n$ (considered with the Euclidean norm).

(5.2.5) Prove Corollary 5.2.4.

(5.2.6) Let $\mathcal{X}, \mathcal{Y}$ be finite-dimensional vector spaces, and $T : \mathcal{X} \to \mathcal{Y}$ linear. Show that

$$\dim\ker T + \dim\operatorname{ran} T = \dim\mathcal{X}.$$

This equality is usually called the **Rank-Nullity Theorem**.

## 5.3. Direct Sums and Quotient spaces

In this section we consider some basic techniques that allow us to create one Banach space out of several, and to decompose a Banach space in terms of subspaces.

**5.3.1. Direct Sums.** Informally, two subspaces $M, N$ of the vector space $\mathcal{X}$ form a direct sum if every element of $\mathcal{X}$ can be written as a sum of an element of $M$ and an element of $N$, and this can done in a unique way. The uniqueness is equivalent with the condition $M \cap N = \{0\}$ (Exercise 5.3.1). In the case of Hilbert spaces we had several instances of this situation; concretely when $\mathcal{K} \subset \mathcal{H}$ is a closed subspace and we write $\mathcal{H} = \mathcal{K} + \mathcal{K}^\perp$. In any finite dimensional vector space $\mathcal{X}$ we can take a subspace $M$, get a basis for $M$, and complete it to a basis for $\mathcal{X}$; the elements of the basis not belonging to $M$ will span a subspace $N$, and $X = M + N$ as a direct sum; such $N$ is not unique.

The notion described above is what is often call an **internal direct sum**. On the other hand, the following would be the **external direct sum**:

**5.3.1. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. The **topological direct sum** $\mathcal{X} \oplus_T \mathcal{Y}$ is the normed vector space $\mathcal{X} \times \mathcal{Y}$, with the pointwise operations and the norm $\|(x, y)\| = \|x\| + \|y\|$.*

For the reader seeing this notion for the first time, let us emphasize that the definition of topological direct sum makes no assumption on what $\mathcal{X}$ and $\mathcal{Y}$ are. For instance there is no issue in considering something like $\mathbb{C}^2 \oplus_T C[0,1]$. We also mention that there are many possible choices for the norm, all equivalent. For instance the norms $\|(x, y)\|_1 = \|x\| + \|y\|$, and $\|(x, y)\|_\infty = \max\{\|x\|, \|y\|\}$, and $\|(x, y)\|_2 = (\|x\|^2 + \|y\|^2)^{1/2}$, are all equivalent with the same proof as before (Exercise 5.3.2).

**5.3.2. Definition.**

*Let $\mathcal{X}$ be a normed vector space, and $M, N \subset \mathcal{X}$ closed subspaces. If $\mathcal{X} = M + N$ and $M \cap N = \{0\}$, we say that $\mathcal{X}$ is the **algebraic direct sum** of $M$ and $N$, and we denote this by $\mathcal{X} = M \oplus N$. We say that $M$ is **complemented** in $\mathcal{X}$, and $N$ is its (algebraic) **complement**.*

*We say that $\mathcal{X}$ is the **topological direct sum** of $M$ and $N$ if $\mathcal{X}$ is **topologically isomorphic** to $M \oplus_T N$. By topologically isomorphic we mean that the map $\gamma : \mathcal{X} \to M \oplus_T N$ given by $\gamma(m + n) = (m, n)$ for $m \in M$, $n \in N$ is linear, bijective, continuous, and with $\gamma^{-1}$ continuous,*

*5.3.3. Remark*. In the case of Banach spaces, the continuity requirement for the inverse will be seen to be unnecessary later, via Theorem 6.3.6. We will also see, with more tools at our disposal (Proposition 6.3.9) that every algebraic direct sum of closed spaces in a Banach space is topological.

*5.3.4. Example*. Any finite-dimensional subspace of any normed space is complemented (Exercise 5.7.2), but the closed subspace $c_0 \subset \ell^\infty(\mathbb{N})$ is not complemented, as we will see soon in Proposition 5.3.6. This is interesting because if one begins with a subspace $M$, it is always possible to choose a (Hamel) basis, extend it to a basis of the full space, and form $N$ as the span of the new elements in the basis. The problem is that even if $M$ is closed, there is no reason for $N$ to be closed (otherwise, $M$ would be complemented, which it need not be). What this means is that we may have $\mathcal{X} = M + N$ with $M \cap N = \{0\}$, but $N$ not closed; when $M$ is not complemented, we have $M \cap \overline{N} \neq \{0\}$. In such situation the projection $P$ in Proposition 5.3.5 will fail to be continuous.

**5.3.5. Proposition.**

*Let $\mathcal{X}$ be a normed space and $M, N \subset \mathcal{X}$ closed subspaces. The following statements are equivalent, where the isomorphism always maps $m+n \longmapsto (m,n)$ for all $m \in M$ and $n \in N$:*

(i) $\mathcal{X} \simeq M \oplus_T N$;

(ii) $\mathcal{X} \simeq M \times N$ *with the norm* $\|(x,y)\| = \max\{\|x\|, \|y\|\}$;

(iii) $\mathcal{X} \simeq M \times N$ *with the norm* $\|(x,y)\| = \sqrt{\|x\|^2 + \|y\|^2}$;

(iv) *There exists a bounded (that is, continuous) projection $P$ onto $M$, with $(I-P)\mathcal{X} = N$; that is, a linear bounded map $P : \mathcal{X} \to \mathcal{X}$ such that $P^2 = P$ and $Px \in M$, $(I-P)x \in N$ for all $x \in \mathcal{X}$.*

*Proof.* All we need to show the equivalence of the first three statements is that $M \oplus_T N \simeq M \oplus_\infty N \simeq M \oplus_2 N$. The isomorphism, of course, will be the identity map. So our claim follows from

$$\|x\| + \|y\| \leq 2\max\{\|x\|, \|y\|\} \leq 2\sqrt{\|x\|^2 + \|y\|^2} \leq 2(\|x\| + \|y\|).$$

(i) $\Longrightarrow$ (iv) Assume now that $\mathcal{X} \simeq M \oplus_T N$ via an isomorphism $\pi$ with $\pi(m) = (m, 0)$ for all $m \in M$. Define $P : \mathcal{X} \to \mathcal{X}$ by $P(x) = \pi_1 \circ \pi(x)$, where $\pi_1$ is the projection onto the first component. By hypothesis $\pi$ is continuous and thus so is $P$. Since $Px \in M$ for all $x \in \mathcal{X}$, we have $P^2x = P(Px) = \pi_1(\pi(Px)) = \pi_1(Px, 0) = Px$, so $P^2 = P$. Given $x \in \mathcal{X}$, we have $\pi(x) = (m, n)$ with $m \in M$ and $n \in N$, so $x = m + n$. Then $(I-P)x = n$.

(iv) $\Longrightarrow$ (i) Since $P$ is continuous we can define $\pi : \mathcal{X} \to M \oplus_T N$ by $\pi(x) = (Px, (I-P)x)$. This is clearly linear. As $P$ is continuous, so is $I - P$. If $x_n \to x$, we have $Px_n \to Px$ and $(I-P)x_n \to (I-P)x$, making $\pi$ continuous. Given $y \in M$, $z \in N$, $\pi(y+z) = (y, z)$, so $\pi$ is onto. And as $x = Px + (I-P)x$, we see that $\pi$ is also one-to-one. It remains to see that $\pi^{-1}$ is continuous. Suppose that $(y_n, z_n) \to (y, z)$. This means that $\|y_n - y\| + \|z_n - z\| \to 0$. Then $\|y_n - y\| \to 0$ and $\|z_n - z\| \to 0$. So

$$\pi^{-1}(y_n, z_n) = y_n + z_n \to y + z = \pi^{-1}(y, z)$$

and $\pi^{-1}$ is continuous. ∎

Proposition 5.3.6 and Remark 5.3.7 below can be safely skipped if this is a first read. They talk about complemented subspaces but, while the arguments are mostly elementary, they are not trivial. These two results will only be relevant in the next chapter.

**5.3.6. *Proposition.***

*The closed subspace $c_0$ is not complemented in $\ell^\infty(\mathbb{N})$.*

*Proof.* This is the scheme of the argument.

(1) Given a countable infinite set $S$, there exists an uncountable family $\{S_r\}_{r\in R}$ of subsets of $S$ such that $|R| = |\mathbb{R}|$, $|S_r| = |\mathbb{N}|$ for all $r$, and $|S_r \cap S_t| < \infty$ whenever $r \neq t$.

(2) Let $P : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ be bounded with $P|_{c_0} = 0$. Then there exists $A \subset \mathbb{N}$, infinite, with $Px = 0$ for all $x$ supported on $A$.

(3) We conclude that $c_0$ is not complemented in $\ell^\infty(\mathbb{N})$.

So let us develop the three steps.

(1) We may identify $S$ with $\mathbb{Q}\cap[0,1]$. Let $R = [0,1]\setminus S$. For each $r \in R$, choose a sequence $\{a_{r,n}\} \subset S$ with $a_{r,n} \to r$; let $A_r = \{a_{r,n} : \; n \in \mathbb{N}\}$. Each $A_r$ is infinite because $r$ is irrational; and $A_r \cap A_t$ is finite if $r \neq t$ (an infinite sequence in the intersection would converge to both $r$ and $t$).

(2) For $S = \mathbb{N}$, consider an uncountable family $\{A_r\}_{r\in R}$ with finite intersections as above. If there exists $r$ such that $Px = 0$ for all $x$ supported in $A_r$, we are done. So assume that, for each $r$, there exists $x_r \in \ell^\infty(\mathbb{N})$, supported in $A_r$, with $Px_r \neq 0$; in particular, $x_r \notin c_0$. By normalizing, we may assume that $\|x_r\| = 1$ for all $r$. Let $R_n = \{r \in R : (Px_r)(n) \neq 0\}$. Since $\bigcup_n R_n = R$, there exists $n$ such that $R_n$ is uncountable. As $R_n = \bigcup_{k\in\mathbb{N}}\{r \in R : \; |(Px_r)(n)| > \frac{1}{k}\}$, there exists $k \in \mathbb{N}$ such that $R_{n,k} = \{r \in R : \; |(Px_r)(n)| > \frac{1}{k}\}$ is infinite (uncountable, but all we need is that it is infinite).

For any finite $K \subset R_{n,k}$, define

$$z = \sum_{r\in K} \exp(-i \arg[(Px_r)(n)])\, x_r \in \ell^\infty(\mathbb{N}).$$

Then

$$(Pz)(n) = \sum_{r\in K} \exp(-i \arg[(Px_r)(n)])\, (Px_r)(n) = \sum_{r\in K} |(Px_r)(n)| \geq \frac{|K|}{k}.$$

Recall that $x_r$ is supported in $A_r$. As any intersection $A_r \cap A_t$ is finite if $t \neq r$, and $K$ is finite, the set $J \subset \mathbb{N}$ of indices $m$ where more than one $x_r(m)$ is nonzero, is finite. Namely, since $x_r(m) \neq 0$

if and only if $m \in A_r$,

$$J = \{m \in \mathbb{N} : \text{ there exist } r_1, r_2 \in K \text{ with } x_{r_1}(m)x_{r_2}(m) \neq 0\}$$
$$= \bigcup_{r_1, r_2 \in K, r_1 \neq r_2} A_{r_1} \cap A_{r_2}$$

is finite.

This allows us to write $z = w + v$, where $w = z|_J$, $v = z|_{\mathbb{N}\setminus J}$. As $|v(m)| = |x_r(m)|$ for some $r$ whenever $v(m) \neq 0$, we have $\|v\| \leq 1$. And as $w$ is finitely supported, it is in $c_0$; so $Pw = 0$. Then

$$\frac{|K|}{k} \leq (Pz)(n) \leq \|Pz\| = \|P(w + v)\| = \|Pv\| \leq \|P\|.$$

This gives us $|K| < k\,\|P\|$, with $k$ fixed and $K$ any finite subset of the infinite set $R_{n,k}$; so we got a contradiction. We have shown that there exists $r$ such that $Px = 0$ for all $x$ supported in $A_r$.

(3) Assume $c_0$ is complemented; this means that $\ell^\infty(\mathbb{N}) = M \oplus c_0$, with $M$ a closed subspace. This allows us to define the bounded projection $P_M$ onto $M$. So $P_M$ is zero on $c_0$ and the above part applies: there exists $A \subset \mathbb{N}$, infinite, with $P_M x = 0$ on all $x$ supported in $A$. This implies that all $x$ supported in $A$ are in $c_0$, a contradiction. $\square$

*5.3.7. Remark*. While $c_0$ is not complemented in $\ell^\infty(\mathbb{N})$, it is complemented in $c$; in fact, in infinitely many ways. Concretely, if $\gamma : c \to \mathbb{C}$ is the limit functional, that is $\gamma(x) = \lim_n x_n$, given any $z \in c$ with $\gamma(z) = 1$ we can define $P_z : c \to c_0$ by $P_z x = x - \gamma(x)\, z$ (note that $c_0 = \ker \gamma$). Any projection $c \to c_0$ is of this form, as we can recover $z$ as $1 - P1$. Namely, given a projection $P : c \to c_0$, since $x \in c$ can be written as $[x - \gamma(x)1] + \gamma(x)1$ and the term in brackets is in $c_0$, we have $Px = x - \gamma(x)1 + \gamma(x)P1 = x - \gamma(x)(1 - P1)$. This implies that $P$ is bounded, for $\|Px\| \leq \|x\|\,(1 + \|1 - P1\|)$. It is also not too hard to show that we always have $\|P\| \geq 2$. Namely, still with $z = 1 - P1$, if we let $Q_n$ be the projection onto the first $n$ coordinates and take $x = z - 2Q_n z$, then $\|x\| = \|(I - 2Q_n)z\| = \|z\|$ (note that all $I - 2Q_n$ does is to multiply the first $n$ entries of $z$ by $-1$), and $\|Px\| = 2\|Q_n z\|$ (since $Pz = 0$). Given $\varepsilon > 0$ we can take $n$ big enough so that $\|Q_n z\| \geq (1 - \varepsilon)\|z\|$. Then

$$\frac{\|Px\|}{\|x\|} = 2\,\frac{\|Q_n z\|}{\|z\|} \geq 2(1 - \varepsilon).$$

In more generality we can consider direct sums of many spaces, and even infinite direct sums.

**5.3.8. Definition.**

*Given Banach spaces $\mathcal{X}_n$, $n \in \mathbb{N}$, we define*

$$\Big(\bigoplus_{n\in\mathbb{N}} \mathcal{X}_n\Big)_{c_0} = \{g : \mathbb{N} \to \bigcup_n \mathcal{X}_n,\ g(n) \in \mathcal{X}_n \ \text{ for all } n,\ \lim_n \|g(n)\| = 0\}$$

*with pointwise vector operations and the norm* $\|g\|_\infty = \sup\{\|g(n)\| : \ n\}$. *Similarly,*

$$\Big(\bigoplus_{n\in\mathbb{N}} \mathcal{X}_n\Big)_{\ell^\infty} = \{g : \mathbb{N} \to \bigcup_n \mathcal{X}_n,\ g(n) \in \mathcal{X}_n \ \text{ for all } n,\ \sup_n \|g(n)\| < \infty\}$$

*with pointwise vector operations and the norm* $\|g\|_\infty = \sup\{\|g(n)\| : \ n\}$. *And, for* $1 \le p < \infty$,

$$\Big(\bigoplus_{n\in\mathbb{N}} \mathcal{X}_n\Big)_{\ell^p} = \{g : \mathbb{N} \to \bigcup_n \mathcal{X}_n,\ g(n) \in \mathcal{X}_n \ \text{ for all } n,\ \sum_n \|g(n)\|^p < \infty\}$$

*with pointwise vector operations and the norm* $\|g\|_p = \Big(\sum_n \|g(n)\|^p\Big)^{1/p}$.

It is routine to verify that in all three cases the infinite direct sum of Banach spaces results in a Banach space (Exercise 5.3.6).

*5.3.9. Remark .* In all three cases above one can easily replace $\mathbb{N}$ with an arbitrary index set, as long as one defines the limit over an arbitrary set in terms of finite sets as we did with series; namely, $\lim_J f(j) = L$ means that given $\varepsilon > 0$ there exists $F_0 \subset J$, finite, such that for every $j \in J \setminus F_0$ we have $|f(j) - L| < \varepsilon$.

The notation we used for the infinite direct sums is common (it appears for example in [**FHH+11**]) but far from universal. Another fairly standard preference is to use $\bigoplus_n \mathcal{X}_n$ for the $c_0$ sum, and $\prod_n \mathcal{X}_n$ for the $\ell^\infty$ one; in such case the former is call *direct sum* and the latter *direct product*. These are a bit simpler, at the cost of not showing explicitly what they are.

**5.3.2. Quotients.** Given a normed vector space and a closed subspace $M \subset \mathcal{X}$, we can consider the quotient $\mathcal{X}/M$ and the canonical quotient map $\rho : \mathcal{X} \to \mathcal{X}/M$, $\rho v = v + M$. In the case of Hilbert spaces, $M$ is complemented, and $\mathcal{X}/M$ can be easily identified with $M^\perp$ (Exercise 5.3.9); so in this case there is a natural norm in the quotient given by the distance to $M$. We will use this idea to define a norm in a general quotient of a normed vector space:

**5.3.10. Definition.**

*Let $\mathcal{X}$ be a normed space and $M \subset \mathcal{X}$ a closed subspace. The **quotient** of $\mathcal{X}$ by $M$ is the normed space $\mathcal{X}/M$, with the norm*

$$\|v + M\| = \inf\{\|v + w\| : \ w \in M\} \tag{5.6}$$

*and the operations $a(v + M) + (w + M) = (av + w) + M$ for $a \in \mathbb{C}$, $v, w \in \mathcal{X}$. The **codimension** of $M$ is the—possibly infinite—number* $\dim \mathcal{X}/M$.

The fact that $M$ is closed implies that (5.6) is a norm (Exercise 5.3.10), since

$$\|v + M\| = \operatorname{dist}(v, M).$$

If $M$ is not required to be closed, then (5.6) will fail to give a norm. For an extreme case, when $M$ is dense then $\|v + M\| = 0$ for all $v$.

*5.3.11. Examples*. Let $\mathcal{X} = C[0, 1]$ with the norm $\|f\|_\infty = \max\{|f(t)| : \ t \in [0, 1]\}$. Let us consider a few closed subspaces and the corresponding quotients.

(a) Let $M = \{f \in \mathcal{X} : \ f(1) = 0\}$. Then $\mathcal{X}/M = \{c + M : \ c \in \mathbb{C}\}$, and $\|f + M\| = |f(1)|$ (Exercise 5.3.11).

(b) Fix $t_1, \ldots, t_n \in [0, 1]$. Let $M = \{f \in \mathcal{X} : \ f(t_j) = 0, \ j = 1, \ldots, n\}$. Then $\mathcal{X}/M \simeq \mathbb{C}^n$, and $\|f + M\| = \max\{|f(t_j)| : \ j = 1, \ldots, n\}$ (Exercise 5.3.12).

(c) Let $M = \left\{f \in \mathcal{X} : \ \displaystyle\int_0^1 f = 0\right\}$. Then $\mathcal{X}/M \simeq \mathbb{C}$, and $\|f + M\| = \left|\displaystyle\int_0^1 f\right|$ (Exercise 5.3.13).

*5.3.12. Example*. Let $\mathcal{X} = \ell^\infty(\mathbb{N})$ and $M = c_0$. Then the norm on $\mathcal{X}/M$ is given by

$$\|a + c_0\| = \limsup_n |a_n|$$

(Exercise 5.3.14).

Here is the main information we need to know about the quotient map.

***5.3.13. Proposition.***

*Let $\mathcal{X}$ be a normed space and $M \subset \mathcal{X}$ a closed subspace. Then*

*(i) $\|\rho(x)\| \leq \|x\|$ for all $x \in \mathcal{X}$; in particular, $\rho$ is continuous;*

(ii) *If $\mathcal{X}$ is a Banach space, so is $\mathcal{X}/M$;*

(iii) *$\rho$ is an open map.*

*Proof.*

(i) Directly,
$$\|\rho(v)\| = \inf\{\|v+w\| : w \in M\} \leq \|v+0\| = \|v\|.$$
This implies that $\rho$ is continuous: if $v_n \to v$, then
$$\|\rho(v_n) - \rho(v)\| = \|\rho(v_n - v)\| \leq \|v_n - v\| \to 0.$$

(ii) Let $\{v_k + M\}$ be a Cauchy sequence in $\mathcal{X}/M$. By choosing a subsequence if necessary, we may assume that $\|v_k - v_{k+1} + M\| < 2^{-k}$. Working on a subsequence does not hamper the argument, since the convergence of a subsequence of a Cauchy sequence implies the convergence of the whole sequence. Put $w_1 = 0$, and choose $w_2' \in M$ such that
$$\|v_1 - v_2 + w_2'\| \leq \|v_1 - v_2 + M\| + 1/2 < 1/2 + 1/2 = 1;$$
then put $w_2 = -w_2'$. Choose $w_3' \in M$ with
$$\|v_2 - v_3 + w_3'\| \leq \|v_2 - v_3 + M\| + 1/4 < 1/4 + 1/4 = 1/2;$$
and let $w_3 = w_2 - w_3'$. Continuing in the same way, choose $w_{k+1}' \in M$ with
$$\|v_k - v_{k+1} + w_{k+1}'\| \leq \|v_k - v_{k+1} + M\| + 2^{-k} < 2^{-k} + 2^{-k} = 2^{-(k-1)}$$
and let $w_{k+1} = w_k - w_{k+1}'$. Note that
$$\begin{aligned}\|(v_k + w_k) - (v_{k+j} + w_{k+j})\| &\leq \sum_{h=k}^{k+j-1} \|(v_h + w_h) - (v_{h+1} + w_{h+1})\| \\ &= \sum_{h=k}^{k+j-1} \|v_h - v_{h+1} + w_{h+1}'\| \\ &\leq \sum_{h=k}^{k+j-1} 2^{-(h-1)} \leq 2^{-(k-1)},\end{aligned}$$
implying that the sequence $\{v_k + w_k\}$ is Cauchy. As $\mathcal{X}$ is complete, there exists $v \in \mathcal{X}$ such that $v_k + w_k \to v$ in $\mathcal{X}$. Then $v_k + M = \rho(v_k + w_k) \to \rho(v) = v + M$.

(iii) Let $W \subset \mathcal{X}$ be open. We want to show that $\rho(W)$ is open. Let $w \in W$. As $W$ is open, the point $w$ is interior. That means that there exists $r > 0$ such that $\|w - y\| < r$ implies $y \in W$. Recall that $\rho$ is surjective. Now, if $\|\rho(w) - \rho(y)\| < r$, then $\|w - y + M\| < r$. Thus there exists $z \in M$ with $\|w - y + z\| < r$, and hence $y - z \in W$. Then $\rho(y) = \rho(y - z) \in \rho(W)$, showing that the ball of radius $r$ around $\rho(w)$ is in $\rho(W)$. $\square$

### 5.3.3. Exercises.

(5.3.1) Let $\mathcal{X}$ be a vector space and $M, N \subset \mathcal{X}$ subspaces with $\mathcal{X} = M + N$. Show that the following statements are equivalent:

(i) each element $x \in \mathcal{X}$ can be written as $x = y + z$, with $y \in M$ and $z \in N$, in a unique way;

(ii) $M \cap N = \{0\}$.

(5.3.2) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. Show that the norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$—as defined after Definition 5.3.1—are equivalent.

(5.3.3) Let $\mathcal{X}$ be a Banach space, and $M \subset \mathcal{X}$ a subspace. Show that the following statements are equivalent:

(i) $M$ is closed;

(ii) if $\{x_n\} \subset M$ is Cauchy, $x = \lim_n x_n$ exists and $x \in M$.

(5.3.4) Let $\mathcal{X}$ be a normed space (not necessarily complete), $M \subset \mathcal{X}$ a subspace. Show that the following statements are equivalent:

(i) $M$ is closed;

(ii) if $\{x_n\} \subset M$ and $x_n \to x$, then $x \in M$.

(5.3.5) Let $\mathcal{X}$ be a Banach space and $P \in \mathcal{B}(\mathcal{X})$ a projection. Show that $P$ has closed range.

(5.3.6) Show that, given a family $\{\mathcal{X}_j\}_{j \in J}$ of Banach spaces, each of

$$\Big(\bigoplus_{j \in J} \mathcal{X}_j\Big)_{c_0}, \qquad \Big(\bigoplus_{j \in J} \mathcal{X}_j\Big)_{\ell^p},$$

$1 \leq p \leq \infty$, is a Banach space.

(5.3.7) Show that, in the particular case where $\mathcal{X}_n = \mathcal{X}$ for all $n$, and $\bigoplus_{n \in \mathbb{N}} \mathcal{X}$ denoting any of the three kind of sums in Exercise 5.3.6,

(i) $\bigoplus_{n \in \mathbb{N}} \mathcal{X} \simeq \mathcal{X} \oplus \bigoplus_{n \in \mathbb{N}} \mathcal{X}$;

(ii) $\bigoplus_{n \in \mathbb{N}} \mathcal{X} \simeq \bigoplus_{n \in \mathbb{N}} \mathcal{X} \oplus \bigoplus_{n \in \mathbb{N}} \mathcal{X}$.

In all cases the isomorphisms are isometric.

(5.3.8) In the setting of Exercise 2.3.27, show that

$$\left(\bigoplus_j L^2(K,\mu_j)\right)_{\ell^2} \simeq L^2(X,\Sigma,\mu),$$

(5.3.9) If $\mathcal{H}$ is a Hilbert space and $M \subset \mathcal{H}$ is a closed subspace, show that $\mathcal{H}/M$ has a natural Hilbert space structure that makes it a Hilbert space, and that $\mathcal{H}/M$ can be identified with (i.e. it is naturally isomorphic to) $M^\perp$.

(5.3.10) Prove that when $M \subset \mathcal{X}$ is a closed subspace, the quotient norm is a norm *(Hint: think of it as a distance)*.

(5.3.11) Let $\mathcal{X} = C[0,1]$ and $M = \{f \in \mathcal{X} : \ f(1) = 0\}$. Show that $\mathcal{X}/M = \{c + M : \ c \in \mathbb{C}\}$ and $\|f + M\| = |f(1)|$.

(5.3.12) Let $\mathcal{X} = C[0,1]$. Fix $t_1, \ldots, t_n \in [0,1]$. Let $M = \{f \in \mathcal{X} : \ f(t_j) = 0, \ j = 1, \ldots, n\}$. Show that $\mathcal{X}/M \simeq \mathbb{C}^n$ and

$$\|f + M\| = \max\{|f(t_j)| : \ j = 1, \ldots, n\}.$$

(5.3.13) Let $\mathcal{X} = C[0,1]$. Let $M = \left\{f \in \mathcal{X} : \ \int_0^1 f = 0\right\}$. Show that $\mathcal{X}/M \simeq \mathbb{C}$ and that $\|f + M\| = \left|\int_0^1 f\right|$.

(5.3.14) Let $\mathcal{X} = \ell^\infty(\mathbb{N})$ and $M = c_0$. Show that the norm on $\mathcal{X}/M$ is given by

$$\|a + c_0\| = \limsup_n |a_n|, \qquad a \in \ell^\infty(\mathbb{N}).$$

## 5.4. Locally Convex Spaces

Not all notions of convergence that appear naturally on vector spaces are given by a norm. For instance we could be dealing with a vector space that is a metric space where the metric does not come from a norm; or even with a topology that does not come from a metric. Consider the following example.

*5.4.1. Example*. Let $B[0,1]$ (bounded Borel functions) with the topology given by pointwise convergence. This convergence cannot be given by a norm. One way to see that this is the case is to show a convergent net with no convergent subsequence. Let $\mathcal{F} = \{F \subset [0,1] : \ \text{finite}\}$, ordered by inclusion and consider the net $\{1_F\}$. It is clear that $1_F \to 1$: for any $t \in [0,1]$, consider all those $F \in \mathcal{F}$ with $F \geq \{t\}$ (that is, $t \in F$): we have $1_F(t) = 1$, so the limit is 1. Now for any subsequence $\{1_{F_n}\}$, the union $\bigcup_n F_n$ is countable, and for any

$t \in [0,1] \setminus \bigcup_n F_n$ we have $1_{F_n}(t) = 0$ for all $n$; so $\{1_{F_n}\}$ does not converge to 1. Thus the convergence cannot be given by a norm (not even by a metric!), since in a metric space choosing a convergent subsequence of a convergent net is always possible.

The theory itself will push us very soon to consider cases as above. Thus the goal of this section is to consider general topological vector spaces, and specialize to the kind we will need later down the road.

**5.4.2. Definition.**

*A **topological vector space** (TVS) is a vector space $\mathcal{X}$ with a topology such that*

*(i) finite subsets of $\mathcal{X}$ are closed, and*

*(ii) the vector space operations (that is, addition and multiplication by scalars) are jointly continuous.*

These apparently naive assumptions in the definition of a topological space have nonetheless far reaching consequences. Condition (i) is equivalent to require $\mathcal{X}$ to be $T_1$. In principle this means that $\mathcal{X}$ is not necessarily Hausdorff by definition (Exercise 5.4.2); but we will soon see that in fact every TVS is Hausdorff (Corollary 5.4.7).

We also mention that one could of course consider vector spaces that only satisfy (ii). But from a practical point of view this would allow very pathological cases and most theorems would need additional hypotheses to weed out the bad cases. Even then, we will have to do a bit of this, as some topological vector spaces are nicer than others.

Throughout this section, $\mathcal{X}$ will stand for a topological vector space.

Let $x \in \mathcal{X}$ and $V$ a neighbourhood of 0. Then, because addition is assumed continuous, $x+V$ is an open neighbourhood of $x$, and every open neighbourhood of $x$ arises this way (Exercise 5.4.3); we say that the topology of $\mathcal{X}$ is **linear**. An equivalent formulation for the linearity of the topology is that $x_j \to x$ if and only if $x_j - x \to 0$.

An observation that will be used repeatedly is that for any subset $S \subset \mathcal{X}$, $S + V$ is open, because $S + V = \bigcup_{s\in S}(s + V)$.

In a metric space, we often use special open sets for most of our dealings: namely, balls. Here, we will need some analog(s) of balls; so we begin by characterizing properties that make balls nice. In general, not all will occur simultaneously.

**5.4.3. Definition.**

*Let $M \subset \mathcal{X}$ be a subset. We say that $M$ is*

*(i)* ***balanced****, if for every $c \in \mathbb{C}$ with $|c| \leq 1$, $cM \subset M$;*

*(ii)* ***absorbing****, if for every $x \in \mathcal{X}$ there exists $t > 0$ with $tx \in M$.*

*(iii)* ***convex****, if for every $x, y \in M$, $t \in [0,1]$, $tx + (1-t)y \in M$.*

We remark that every open neighbourhood of 0 in a topological vector space is absorbing, and that any nonzero multiple of an open set is open (Exercise 5.4.4). An immediate consequence of the definition is that if $M$ is balanced then $cM = M$ for all $c \in \mathbb{T}$ (Exercise 5.4.1).

***5.4.4. Lemma.***

*Let $\mathcal{X}$ be a topological vector space. If $V \subset \mathcal{X}$ is an open neighbourhood of $0$, then there is a balanced open neighbourhood $W$ of $0$, such that $W + W \subset V$.*

*Proof.* By definition, scalar multiplication is continuous, i.e. the map $(c, x) \mapsto cx$ is continuous from $\mathbb{C} \times \mathcal{X} \to \mathcal{X}$. So we can find a neighbourhood $W_0$ of 0 and $\delta > 0$ such that $cW_0 \subset V$ whenever $|c| < \delta$. Then $\bigcup_{|c|<\delta} cW_0$ is an open balanced subset of $V$, containing 0.

From the fact that addition is continuous and $0 + 0 = 0$, given the neighbourhood $V$ we can get neighbourhoods $W_1$, $W_2$ of 0, with $W_1 + W_2 \subset V$. By the first paragraph, we may assume that $W_1$, $W_2$ are balanced by shrinking them if necessary. Now put $W = W_1 \cap W_2$. Then $W$ is balanced, $W + W$ is open (Exercise 5.4.5) and $W + W \subset V$. □

Convexity is another property from balls that we will often be useful; this will be particularly noticeable in Chapter 7.

***5.4.5. Lemma.***

*Every convex open neighbourhood of $0$ contains a balanced convex open neighbourhood of $0$.*

*Proof.* Let $U$ be a convex open neighbourhood of 0. By Lemma 5.4.4, there exists $W$ open, balanced, with $0 \in W \subset U$. Since $W$ is balanced, $cW = W$ for all $c \in \mathbb{C}$ with $|c| = 1$ (Exercise 5.4.1). Let $V' = \bigcap_{|c|=1} cU$. Then $W \subset V'$, so $V'$ has nonempty interior and 0 is an interior point of $V'$. Let $V$ be the interior of $V'$. Since $V'$ is an intersection of convex sets, it is convex; since

$V$ is the interior of a convex set, it is convex (Exercise 5.4.6). It remains to show that $V$ is balanced. For that, we will show that $V'$ is balanced, and then the result will follow from Exercise 5.4.7. Since $U$ is convex and $0 \in U$, $tU \subset U$ for all $t \in [0,1]$. For $d \in \mathbb{C}$ with $|d| \leq 1$, write $d = te^{i\theta}$, $t \in [0,1]$. Then

$$dV' = te^{i\theta} \bigcap_{|c|=1} cU = \bigcap_{|c|=1} te^{i\theta} cU = \bigcap_{|c|=1} ctU \subset \bigcap_{|c|=1} cU = V'.$$

So $V'$ is balanced, and so is $V$. ☐

The big consequence of the assumption that finite sets are closed is the following much stronger condition:

**5.4.6. Theorem.**

*Let $\mathcal{X}$ be a topological vector space, $K, C \subset \mathcal{X}$, with $K$ compact and $C$ closed, and $K \cap C = \varnothing$. Then there exists an open neighbourhood $V$ of $0$ such that $(K+V) \cap (C+V) = \varnothing$.*

*Proof.* Let $x \in K$. Then, since the complement of $C$ is open, there exists a neighbourhood $V_x'$ of zero with $(x+V_x') \cap C = \varnothing$. By Lemma 5.4.4, there exists $V_x''$, balanced, with $V_x'' + V_x'' \subset V_x'$. Applying the Lemma once more, we get balanced $V_x$ with $V_x + V_x \subset V_x''$. Then $x + (V_x + V_x + V_x) \cap C = \varnothing$. Because $V_x$ is balanced, $V_x = -V_x$, and so we can deduce that $(x+V_x+V_x) \cap (C+V_x) = \varnothing$. The compactness of $K$ then gives us points $x_1, \ldots, x_k \in K$ with

$$K \subset (x_1 + V_{x_1}) \cup \cdots \cup (x_k + V_{x_k}), \qquad (x_j + V_{x_j} + V_{x_j}) \cap (C + V_{x_j}) = \varnothing. \tag{5.7}$$

If we let $V = V_{x_1} \cap \cdots \cap V_{x_k}$, then

$$K + V \subset \bigcup_{j=1}^{k} (x_j + V_{x_j} + V) \subset \bigcup_{j=1} (x_j + V_{x_j} + V_{x_j})$$

and the last union is in the complement of $C + V$ by (5.7). ☐

In particular:

**5.4.7. Corollary.**

*A topological vector space is Hausdorff.*

*Proof.* Given $x, y \in \mathcal{X}$, apply Theorem 5.4.6 to $K = \{x\}$ and $C = \{y\}$ ($C$ is closed by definition of topological vector space). ☐

In a normed space we can always write $x = \|x\|\, x_1$, with $\|x_1\| = 1$; that is, $x_1$ is in the closed unit ball. If the ball $B$ is given, we can recover $\|x\|$ by noticing that it is characterized as the smallest number $c$ such that $\frac{1}{c}\, x \in B$. That is,

$$\|x\| = \inf\left\{c > 0 :\ \frac{1}{c}\, x \in B\right\} = \inf\left\{\frac{1}{t} :\ t > 0,\ tx \in B\right\}.$$

This should make the following definition natural. Do not forget that now $\mathcal{X}$ does not necessarily come with a norm!

**5.4.8. Definition.**

*Let $M \subset \mathcal{X}$ be an absorbing subset. The **Minkowski functional** of $M$ is*

$$\mu_M(x) = \inf\left\{\frac{1}{t}\ :\ t > 0 \ \text{ and } tx \in M\right\}$$

*A **sublinear function** on $\mathcal{X}$ is a function $q : \mathcal{X} \to \mathbb{R}$ such that*

(i) $q(x + y) \leq q(x) + q(y)$ (***subadditive***);

(ii) $q(tx) = tq(x)$ *for all* $t > 0$ (***homogeneous***).

Note that the definition implies $q(0) = 0$ (from $q(0) = q(2 \times 0) = 2q(0)$).

The usefulness of the Minkowski functional comes from the following result. As suggested above, in spirit we are taking $M$ as "a unit ball"—emphasis on "a"—when we define the Minkowski functional. If $M$ shares the key properties a ball has, this works fairly well:

***5.4.9. Proposition.***

*Let $\mathcal{X}$ be a topological vector space and $M \subset \mathcal{X}$ an open convex neighbourhood of $0$. Then $\mu_M$ is sublinear and*

$$M = \{x \in \mathcal{X} :\ \mu_M(x) < 1\}.$$

*If moreover $M$ is balanced, then $\mu_M$ is a seminorm.*

*Proof.* The fact that $M$ is absorbing guarantees that $\mu_M$ is well-defined in all of $\mathcal{X}$. For $s > 0$,

$$\mu_M(sx) = \inf\left\{\frac{1}{t}\ :\ tsx \in M\right\} = \inf\left\{\frac{s}{t}\ :\ tx \in M\right\} = s\,\mu_M(x)$$

(for the second equality, think of doing a substitution $u = st$).

When $M$ is balanced, the definition of balanced implies that if $c \in \mathbb{C}$ with $|c| = 1$, then $cM = M$ (Exercise 5.4.1). A consequence of this is that $\mu_M(cx) = \mu_M(x)$ when $|c| = 1$. Then, for any $c \in \mathbb{C}$, write $c = |c|e^{i\theta}$ and then

$$\mu_M(cx) = \mu_M(|c|e^{i\theta}x) = |c|\mu_M(e^{i\theta}x) = |c|\mu_M(x).$$

Fix $x, y \in \mathcal{X}$ and $\varepsilon > 0$. Then we can find $r, s > 0$ with $r < \mu_M(x) + \varepsilon/2$, $s < \mu_M(y) + \varepsilon/2$, and $x/r \in M$, $y/s \in M$. Using that $M$ is convex,

$$\frac{1}{r+s}(x+y) = \frac{r}{r+s}\frac{x}{r} + \frac{s}{r+s}\frac{y}{s} \in M.$$

We conclude that $\mu_M(x+y) \leq r+s < \mu_M(x) + \mu_M(y) + \varepsilon$. With $\varepsilon$ arbitrary, this proves that the triangle inequality is satisfied.

Finally, since the map $t \mapsto tx$ is continuous everywhere (and hence at $t = 1$) and $M$ is open, for $x \in M$ there exists $t > 1$ with $tx \in M$: then $\mu_M(x) < 1$. Conversely, if $\mu_M(x) < 1$, there exists $t > 1$ with $tx \in M$; as $M$ is convex and $0 \in M$, $x = \frac{1}{t}(tx) + \left(1 - \frac{1}{t}\right) 0 \in M$. So $M = \{x \in \mathcal{X} : \mu_M(x) < 1\}$. □

The strength of the tools we have developed is that we are now in position to characterize locally convex topological vector spaces. These will appear prominently in Chapter 7 and beyond.

**5.4.10. Definition.**

*A topological vector space $\mathcal{X}$ is* ***locally convex*** *if every point has a local base made of convex sets.*

*A family $\mathcal{P}$ of seminorms on $\mathcal{X}$* ***separates points*** *if $p(x) = 0$ for all $p \in \mathcal{P}$ implies that $x = 0$.*

Any normed space is locally convex as every point has a base made of balls (and balls from a norm are convex, as can be seen from a straightforward computation). Soon, after we talk about seminorms, we will have other examples available.

A family $\mathcal{P}$ of seminorms induces a linear topology by taking as a local base at 0 the sets

$$V(p_1, \ldots, p_n; \varepsilon) = \{x \in \mathcal{X} : \; p_j(x) < \varepsilon, \; j = 1, \ldots, n\} \tag{5.8}$$

indexed by $n \in \mathbb{N}$, $\varepsilon > 0$, and $p_1, \ldots, p_n \in \mathcal{P}$ (Exercise 5.4.13). The fact that the $p_j$ are seminorms guarantees that each $V(p_1, \ldots, p_n, \varepsilon)$ is convex, so the induced topology is locally convex. Remarkably, the converse also holds:

***5.4.11. Theorem.***

*Let $\mathcal{X}$ be a topological vector space. Then the following statements are equivalent:*

*(i) $\mathcal{X}$ is locally convex;*

*(ii) the topology of $\mathcal{X}$ is given by a family of seminorms that separates points.*

*Proof.* The implication (ii) $\implies$ (i) is immediate—as shown above—so we only prove the other implication.

If $\mathcal{X}$ is locally convex, by Lemma 5.4.5 there is a local base $\mathcal{V}$ at 0 made of balanced convex open sets. By Proposition 5.4.9

$$\mathcal{P} = \{\mu_V : V \in \mathcal{V}\}$$

is a family of seminorms, with

$$V = \{x \in \mathcal{X} : \mu_V(x) < 1\} \tag{5.9}$$

for all $V \in \mathcal{V}$. The family $\mathcal{P}$ clearly separates points, because if $\mu_V(x) = 0$ for all $V$, it means that $x$ lies in every open neighbourhood of 0, and so $x = 0$.

By equation (5.9), the topology induced by $\mathcal{P}$ is finer than the topology of $\mathcal{X}$, since the convex, balanced, open sets generate the topology by Lemma 5.4.5 and the hypothesis that $\mathcal{X}$ is locally convex; and a convex, balanced, open $W \subset \mathcal{X}$ satisfies $W = V(\mu_W; 1)$, in the notation of (5.8). On the other hand, given $\varepsilon > 0$ and $V_1, \ldots, V_n \in \mathcal{V}$, and using Exercise 5.4.11,

$$\begin{aligned} V(\mu_{V_1}, \ldots, \mu_{V_n}; \varepsilon) &= \bigcap_{j=1}^{n} \{x : \mu_{V_j}(x) < \varepsilon\} = \bigcap_{j=1}^{n} \{x : \mu_{\varepsilon V_j}(x) < 1\} \\ &= \bigcap_{j=1}^{n} \varepsilon V_j. \end{aligned}$$

This shows that the topology induced by $\mathcal{P}$ is coarser than the topology of $\mathcal{X}$. So $\mathcal{P}$ determines the topology of $\mathcal{X}$. $\square$

*5.4.12. Remark.* No generality is lost if one fixes $\varepsilon = 1$ in (5.8). Indeed, as positive multiples of seminorms are again seminorms,

$$V(p_1, \ldots, p_n; \varepsilon) = V(\varepsilon^{-1} p_1, \ldots, \varepsilon^{-1} p_n; 1).$$

We mention below some examples of locally convex spaces. Natural examples will arise in Section 5.6 and Chapter 7.

*5.4.13. Example.* As already mentioned, any normed space is a locally convex space, with $\mathcal{P} = \{\| \cdot \|\}$.

*5.4.14. Example.* An example of a not-normed locally convex space was mentioned in Example 5.4.1, where the topology is given by the family of seminorms $\mathcal{P} = \{p_t : t \in [0, 1]\}$, with $p_t(f) = |f(t)|$. We can similarly consider $\mathcal{X} = C(T)$ for $T$ any compact Hausdorff space.

*5.4.15. Example.* Let $\mathcal{H}$ be a Hilbert space and consider the topology given by weak convergence; that is, $\xi_j \xrightarrow{\text{weak}} \xi$ means $\langle \xi_j - \xi, \eta \rangle \to 0$ for all $\eta \in \mathcal{H}$. This is a locally convex space, where the topology is determined by the family of seminorms $\mathcal{P} = \{p_\eta : \ \eta \in \mathcal{H}\}$, where $p_\eta(\xi) = |\langle \xi, \eta \rangle|$. When $\dim \mathcal{H} < \infty$, the weak topology is metrizable; in fact, it agrees with the original topology (Theorem 5.4.16). But the weak topology on an infinite-dimensional Hilbert space is not metrizable, as we will see in Remark 7.1.13. One can check directly that if $\mathcal{H}$ is infinite-dimensional and $\{\xi_j\}$ is an orthonormal basis, then $\xi_j \xrightarrow{\text{weak}} 0$ (Exercise 5.4.10), which guarantees that the weak topology differs from the norm topology on $\mathcal{H}$.

Locally convex topologies play a big role in Section 5.7, in Chapter 7, and beyond. For now we establish the analogue of Theorem 5.2.2.

***5.4.16. Theorem.***

*Let $\mathcal{X}$ be a finite-dimensional topological vector space. Then there exists a single topological vector space topology on $\mathcal{X}$.*

*Proof.* We will show that $\mathcal{X}$ is homeomorphic to $\mathbb{C}^n$ with the usual topology. Let $\{e_1, \ldots, e_n\}$ be a basis of $\mathcal{X}$. Define $\Gamma : \mathbb{C}^n \to \mathcal{X}$ by $\Gamma(c) = \sum_{j=1}^n c_j e_j$; from the fact that $\{e_1, \ldots, e_n\}$ is a basis, it follows that $\Gamma$ is a bijection. As vector space operations are continuous, $\Gamma$ is continuous (here we use crucially that convergence in $\mathbb{C}^n$ is pointwise convergence). The restriction of $\Gamma$ to the unit ball of $\mathbb{C}^n$ is a continuous bijection from a compact space onto its image, a Hausdorff topological space. That's enough to make this restriction a homeomorphism onto its image (Exercise 1.8.38). Now to show that $\Gamma^{-1}$ is continuous, suppose that $x_r \to 0$, where $x_r = \sum_{j=1}^n c_j^{(r)} e_j$. Fix $\varepsilon > 0$, with $\varepsilon < 1$. Since $\varepsilon B_1(0)$ is open in $\overline{B_1(0)} \subset \mathbb{C}^n$ and $\Gamma$ is a homeomorphism when restricted to $\overline{B_1(0)}$, it follows that $\Gamma(\varepsilon B_1(0))$ is open on $\mathcal{X}$, and $0 \in \Gamma(\varepsilon B_1(0))$. So there exists $r_0$ such that $x_r \in \Gamma(\varepsilon B_1(0))$ whenever $r \geq r_0$. Thus $(c_1^{(r)}, \ldots, c_n^{(r)}) = \Gamma^{-1}(x_r) \in \varepsilon B_1(0)$, and we have shown that $(c_1^{(r)}, \ldots, c_n^{(r)}) \underset{r}{\to} 0$, implying that $\Gamma^{-1}$ is continuous. $\square$

*5.4.17. Remark.* We have proven that a finite-dimensional $\mathcal{X}$ is linearly homeomorphic to $\mathbb{C}^n$ with its usual topology, that is given by a norm. Thus, even if $\mathcal{X}$ is given to us with some obscure topological vector space topology, we can define a norm on $\mathcal{X}$ by $\|x\|_{\mathcal{X}} = \|\Gamma^{-1}(x)\|_{\mathbb{C}^n}$. The fact that $\Gamma$ is linear, bijective, and bicontinuous makes $\mathcal{X}$ a Banach space with $\|\cdot\|_{\mathcal{X}}$.

**5.4.1. Direct sums and quotient spaces.** Most of what was done for direct sums and quotients of normed spaces in Section 5.3 still works for locally convex spaces. We will develop the few results we need.

**5.4.18. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be locally convex spaces with their topologies given by the families of seminorms $S_\mathcal{X}$ and $S_\mathcal{Y}$ respectively. The* ***topological direct sum*** *$\mathcal{X} \oplus_T \mathcal{Y}$ is the locally convex space $\mathcal{X} \times \mathcal{Y}$, with the pointwise operations and the family of seminorms*

$$S_{\mathcal{X}\times\mathcal{Y}} = \{p \times q : \ p \in S_\mathcal{X},\ q \in S_\mathcal{Y}\},$$

*where $p \times q$ is the seminorm $(p \times q)(x, y) = p(x) + q(y)$.*

*Given closed subspaces $M, N \subset \mathcal{X}$, we say that $\mathcal{X}$ is the* ***topological direct sum*** *of $M$ and $N$ if $\mathcal{X} = M \oplus N$ and the map $\gamma : \mathcal{X} \to M \oplus_T N$ given by $\gamma(m + n) = (m, n)$ is bijective and bicontinuous.*

The map $\pi_1 : \mathcal{X} \oplus_T \mathcal{Y} \to x$ given by $\pi_1(x, y) = x$ is continuous: it is linear, and if $(x_j, y_j) \to 0$, then in particular $\pi_1(x_j, y_j) = x_j \to 0$.

***5.4.19. Proposition.***

*Let $\mathcal{X}$ be a locally convex space and $M, N \subset \mathcal{X}$ closed subspaces such that $\mathcal{X} = M \oplus N$. Then $\mathcal{X}$ is the topological direct sum of $M$ and $N$ if and only if there exists a continuous projection $P : \mathcal{X} \to M$. That is, $P$ is linear, continuous, and $P^2 = P$.*

*Proof.* If $\mathcal{X}$ is the topological direct sum of $M$ and $N$, let $\gamma : \mathcal{X} \to M \oplus_T N$ be the map $\gamma(m+n) = (m, n)$, which is bijective and bicontinuous by hypothesis. Let $P = \pi_1 \circ \gamma$. Then $P$ is linear and continuous as both $\pi_1$ and $\gamma$ are continuous. Also, for any $m \in M$, $Pm = \pi(m, 0) = m$, so $P^2(m + n) = m = P(m + n)$ and $P^2 = P$.

Conversely, suppose that $P$ exists and is linear and continuous. Let $\gamma(x) = (Px, (I - P)x)$. As in the proof of Proposition 5.3.5, $\gamma$ is linear, bijective, and continuous. It remains to check that $\gamma^{-1}$ is continuous. But if $(m_j, n_j) \to 0$, this means that $m_j \to 0$ and $n_j \to 0$, so $m_j + n_j \to 0$. □

Given a locally convex space $\mathcal{X}$ and $M \subset \mathcal{X}$ closed, we consider the quotient $\mathcal{X}/M$ with the quotient topology

$$\tilde{\mathcal{T}} = \{\tilde{V} \subset \mathcal{X}/M : \ q^{-1}(\tilde{V}) \text{ is open in } \mathcal{X}\},$$

where $q : \mathcal{X} \to \mathcal{X}/M$ is the quotient map. We will only consider the case where the subspace $M$ is closed; otherwise we can get things like when $M$ is dense, and the only open sets in $\mathcal{X}/M$ are $\varnothing$ and $\mathcal{X}/M$.

**5.4.20. Lemma.**

*Let $\mathcal{X}$ be a locally convex space and $M \subset \mathcal{X}$ a closed subspace. Then the quotient map $q$ is both continuous and open.*

*Proof.* We have by definition of $\tilde{\mathcal{T}}$ that $q^{-1}(\tilde{V})$ is open for all open $\tilde{V}$, so $q$ is continuous. Conversely, if $W \subset \mathcal{X}$ is open, then $q^{-1}(q(W)) = W + M = \bigcup_{m \in M}(m + W)$ is open in $\mathcal{X}$, so $q(W)$ is open. $\square$

**5.4.21. Proposition.**

*Let $\mathcal{X}$ be a locally convex space with its topology given by the family of seminorms $S_{\mathcal{X}}$ and $M \subset \mathcal{X}$ a closed subspace. Then $\mathcal{X}/M$ is the locally convex space given by the family of seminorms $\tilde{S}_{\mathcal{X}} = \{\tilde{p} : \ p \in S_{\mathcal{X}}\}$, where for $p \in S_{\mathcal{X}}$ the seminorm $\tilde{p}$ is given by*

$$\tilde{p}(x + M) = \inf\{p(x + m) : \ m \in M\}.$$

*Proof.* First of all we need to check that each $\tilde{p}$ is indeed a seminorm and that $\tilde{S}_{\mathcal{X}}$ separates points (Exercise 5.4.15).

So it remains to show that the two topologies $\tilde{\mathcal{T}}$ and $\tilde{T}_{\mathcal{X}}$ agree, where $\tilde{T}_{\mathcal{X}}$ is the topology on $\mathcal{X}/M$ induced by the seminorms $\tilde{S}_{\mathcal{X}}$. Consider $\tilde{V} \in \tilde{\mathcal{T}}$ and $q(x) \in \tilde{V}$. We have that $x \in q^{-1}(\tilde{V})$, which is open. So there exists $W \in \mathcal{T}_{\mathcal{X}}$ with $x \in W \in q^{-1}(\tilde{V})$. As $q$ is an open map, $q(W)$ is open, and $q(x) \in q(W) \subset \tilde{V}$. So $\tilde{V}$ is $\tilde{\mathcal{T}}_{\mathcal{X}}$-open, and thus $\tilde{\mathcal{T}} \subset \tilde{\mathcal{T}}_{\mathcal{X}}$. Conversely, given a neighbourhood of 0 $V(\tilde{p}_1, \ldots, \tilde{p}_s) \subset \mathcal{X}/M$,

$$\begin{aligned}
q^{-1}(V(\tilde{p}_1, \ldots, \tilde{p}_s)) &= \{x \in \mathcal{X} : \ \tilde{p}_j(x + M) < 1, \ j = 1, \ldots, s\} \\
&= \{x \in \mathcal{X} : \exists m_1, \ldots, m_s \ \text{ with } p_j(x + m_j) < 1, \ \forall \ j\} \\
&= \bigcap_{j=1}^{s} \{x : \ \exists m, \ p_j(x + m) < 1\} = \bigcap_{j=1}^{s} \{x - m : \ p_j(x) < 1\} \\
&= \bigcap_{j=1}^{s} V(p_j) + M = \bigcap_{j=1}^{s} \bigcup_{m \in M} V(p_j) + m
\end{aligned}$$

is open in $\mathcal{X}$. So $V(\tilde{p}_1, \ldots, \tilde{p}_s) \in \tilde{\mathcal{T}}$, showing that $\tilde{\mathcal{T}}_{\mathcal{X}} \subset \tilde{\mathcal{T}}$, and hence $\tilde{\mathcal{T}} = \tilde{\mathcal{T}}_{\mathcal{X}}$. $\square$

When $\mathcal{X}$ is normed, the locally convex quotient agrees with the quotient as a normed space, as we can take the norm as the unique element of a family of seminorms.

A consequence of Proposition 5.4.21 is that if we take another family of seminorms $S'_{\mathcal{X}}$ that generates the same topology on $\mathcal{X}$, then the corresponding family of seminorms $\tilde{S}'_{\mathcal{X}}$ generates the same topology on $\mathcal{X}/M$.

**5.4.2. Exercises.**

(5.4.1) Let $\mathcal{X}$ be a topological vector space and $M \subset \mathcal{X}$ balanced. Show that if $c \in \mathbb{T}$ then $cM = M$.

(5.4.2) Let $\mathcal{X}$ be a topological space. Show that the following statements are equivalent:

(i) singletons are closed;

(ii) finite sets are closed;

(iii) $\mathcal{X}$ is $T_1$: namely, given distinct $x, y \in \mathcal{X}$, there exists an open set $V$ such that $x \in V$, $y \notin V$.

(5.4.3) Show that in a topological vector space, all open neighbourhoods around a point $x$ are given by translates by $x$ of neighbourhoods of 0.

(5.4.4) Show that any open neighbourhood of 0 in a topological vector space is absorbing, and that any nonzero multiple of an open set is open.

(5.4.5) Let $\mathcal{X}$ be a TVS and $V, W \subset \mathcal{X}$ with $V$ open. Show that $V + W$ is open.

(5.4.6) Prove that in a topological vector space the interior and the closure of a convex set are convex.

(5.4.7) Prove that, in a topological vector space, the closure of a balanced set is balanced; and if the interior contains 0, then the interior is balanced.

(5.4.8) Give examples in $\mathbb{C}^2$, with the usual topology, of open neighbourhoods of 0 that are:

(i) balanced but not convex;

(ii) convex but not balanced.

In each case, does a local basis at 0 for the topology exist where all sets are like that? Are the same examples possible in $\mathbb{C}$?

(5.4.9) Fill the details in Example 5.4.14, i.e. show that the topology induced by the seminorms agrees with the topology of pointwise-convergence.

(5.4.10) Show that if $\mathcal{H}$ is an infinite-dimensional Hilbert space and $\{\xi_j\}$ is an orthonormal basis, then $\xi_j \xrightarrow{\text{weak}} 0$. *(We defined weak convergence on Example 5.4.13)*

(5.4.11) Let $V \subset \mathcal{X}$ be an open, balanced, convex, set with $0 \in V$. Let $\varepsilon > 0$. Show that

$$\{x : \ \mu_V(x) < \varepsilon\} = \{x : \ \mu_{\varepsilon V}(x) < 1\} = \varepsilon\, V.$$

(5.4.12) Let $\mathcal{X}$ be a TVS and $M \subset \mathcal{X}$ a convex, open, neighbourhood of 0. Show that $\mu_M$ is continuous.

(5.4.13) Let $\mathcal{X}$ be a vector space and $\mathcal{P}$ a family of seminorms that separates points. Show that the sets

$$V_x(p_1, \ldots, p_n, \varepsilon), \qquad x \in \mathcal{X},\ p_1, \ldots, p_n \in \mathcal{P},\ \varepsilon > 0,$$

where

$$V_x(p_1, \ldots, p_n, \varepsilon) = \{x' \in \mathcal{X} :\ p_k(x' - x) < \varepsilon,\ k = 1, \ldots, n\}.$$

form a base for a topology.

(5.4.14) Let $\mathcal{X}, \mathcal{Y}$ be locally convex spaces. Show that $(x_j, y_j) \to (x, y)$ on $\mathcal{X} \oplus_T \mathcal{Y}$ if and only if $x_j \to x$ and $y_j \to y$.

(5.4.15) Show that the family $\tilde{S}_{\mathcal{X}}$ from Proposition 5.4.21 is indeed a family of seminorms that separates points.

## 5.5. The Dual

Let $\mathcal{X}$ be a TVS, and let $\varphi : \mathcal{X} \to \mathbb{C}$ be linear. The usual terminology is that $\varphi$ is a **functional**, or more specifically a **linear functional**.

*5.5.1. Examples*. Here are some very well-known examples of familiar linear functionals.

(a) Let $\mathcal{X} = C[0,1]$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(f) = f(0)$.

(b) Let $\mathcal{X} = C[0,1]$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(f) = \frac{f(0)+f(1)}{2}$.

(c) Let $\mathcal{X} = \mathbb{C}[x]$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(f) = f'(1)$.

(d) Let $\mathcal{X} = C[0,1]$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(f) = \int_0^1 f$.

(e) Let $\mathcal{X} = M_n(\mathbb{C})$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(A) = A_{21}$.

(f) Let $\mathcal{X} = M_n(\mathbb{C})$, and $\varphi : \mathcal{X} \to \mathbb{C}$ the map $\varphi(A) = \mathrm{Tr}(A)$.

We can get an unending supply of other functionals from these, since linear combinations of linear functionals are again linear functionals.

First, a short technical result.

**5.5.2. Proposition.**

*Let $K$ be a (not necessarily closed) subspace of $\mathcal{X}$. Then the following statements are equivalent:*

*(i)* *$K$ is the kernel of a nonzero linear functional.*

*(ii)* *$\mathcal{X}/K$ is one-dimensional;*

*Proof.* Suppose that $K = \ker\varphi$ for some nonzero linear functional $\varphi$. Fix $x \in \mathcal{X} \setminus K$. Then $\varphi(x) \neq 0$. Given $y \in \mathcal{X} \setminus K$, let $c = \varphi(y)/\varphi(x)$; then $\varphi(y - cx) = 0$, so that $y - cx \in K$. This means that $y + K = c(x + K)$ and this can be done for any $y$, showing that $\mathcal{X}/K$ is one-dimensional.

Conversely, if $\mathcal{X}/K$ is one-dimensional, there exists $x \in \mathcal{X} \setminus K$. Since $\mathcal{X}/K$ is one-dimensional, for any $y \in \mathcal{X}$ there exists a unique scalar $c_y$ such that $y + K = c_y x + K$. Then $\varphi(y) = c_y$ is a linear functional with $K = \ker\varphi$ (Exercise 5.5.4). □

**5.5.3. Definition.**

*A subspace satisfying the conditions in Proposition 5.5.2 is called a* ***hyperplane****.*

A natural question in infinite dimensional spaces is when a functional is continuous. We discussed this, in the context of Hilbert spaces, in Chapter 4. If one looks carefully, the proof of Proposition 4.5.1 only used the fact that $H$ was a normed space, and so it applies to any normed space. We state this fact below:

**5.5.4. Proposition.**

*Let $\mathcal{X}$ be a normed vector space, and let $\varphi : \mathcal{X} \to \mathbb{C}$ be linear. Then the following statements are equivalent:*

*(i)* *$\varphi$ is continuous;*

*(ii)* *$\varphi$ is continuous at $0$;*

*(iii)* *there exists $x \in \mathcal{X}$ such that $\varphi$ is continuous at $x$;*

*(iv)* *there exists $r > 0$ such that, for every $x \in \mathcal{X}$, $|\varphi(x)| \leq r\,\|x\|$.*

*Proof.* Exercise 5.5.5. □

As is the case with functionals over a Hilbert space, a linear functional over a normed vector space that satisfies any of the statements in Proposition 5.5.4 is called a **bounded linear functional**.

In Chapter 4, page 331 we showed an example of an unbounded linear functional. Similar to that we can consider $\mathcal{X} = \mathbb{C}[x]$ and $\varphi : \mathbb{C}[x] \to \mathbb{C}$ given by $\varphi(f) = f'(1)$ but now with the norm $\|f\|_\infty = \max\{|f(t)| : \ t \in [0,1]\}$; it will be again unbounded. Indeed, let $f_n(x) = \frac{x^n}{n}$. Then $\|f_n\|_\infty = \frac{1}{n} \to 0$, while $\varphi(f_n) = 1$ for all $n$. An obvious objection to these examples is that $\mathcal{X}$ is not complete. It turns out that it is not possible to explicitly construct an unbounded linear functional on a complete space; the aforementioned example on page 331 requires a Hamel basis, and the axiom that every vector space has a Hamel basis is equivalent to the Axiom of Choice (AC). What this means is that it is possible to come up with an axiomatic framework—without AC—where all linear functionals on a certain space are bounded. But the AC is essential to Functional Analysis (if anything, for the ever present Hahn–Banach Theorem 5.7.5 and its corollaries) so we have to live with this situation—we refer the reader to Section D.2 for a short discussion on how life without AC is not so rosy. Although beyond what we have covered so far, it is worth mentioning that due to the Closed Graph Theorem (6.3.12), a closed everywhere-defined linear operator is bounded. And "naturally-defined" linear operators are closed, so necessarily the common unbounded linear operators—like the example above with the derivative—will be densely-defined and not everywhere-defined.

In light of Proposition 5.5.4 we make the following definition:

**5.5.5. Definition.**

*Let $\varphi$ be a bounded linear functional on the normed space $\mathcal{X}$. The **norm** of $\varphi$ is the number*

$$\|\varphi\| = \inf\{r : |\varphi(x)| \le r\|x\| \text{ for all } x \in \mathcal{X}\}. \tag{5.10}$$

Of course one needs to prove that such assignment is indeed a norm (Exercise 5.5.7). Also, the definition in (5.10) is not very practical. Most commonly, one uses one of the last two characterizations below.

***5.5.6. Proposition.***

*Let $\varphi$ be a bounded linear functional on the normed space $\mathcal{X}$. Then*

*(i)* $\|\varphi\| = \min\{r : |\varphi(x)| \le r\|x\| \text{ for all } x \in \mathcal{X}\}$;

*(ii)* $\|\varphi\| = \sup\{|\varphi(x)| : \ \|x\| = 1\}$;

*(iii)* $\|\varphi\| = \sup\{|\varphi(x)|/\|x\| : \ x \neq 0\}$.

*Moreover,*

$$|\varphi(x)| \le \|\varphi\|\,\|x\|, \qquad x \in \mathcal{X}.$$

*Proof.* Exercise 5.5.6. ☐

We remark that the sup in

$$\|\varphi\| = \sup\{|\varphi(x)| : \ \|x\| = 1\}$$

is not realized in general. For an easy example consider $\mathcal{X} = \{f \in C[0,1] : f(0) = 0\}$ with the supremum norm, and $\varphi(f) = \int_0^1 f$. Then $\|\varphi\| = 1$ but $|\varphi(f)| < 1$ for all $f \in \mathcal{X}$ with $\|f\|_\infty = 1$ (Exercise 5.5.9).

**5.5.7. Definition.**

*The* ***dual space*** *of a normed vector space $\mathcal{X}$ is the space*

$$\mathcal{X}^* = \{\varphi : \mathcal{X} \to \mathbb{C}, \quad \textit{linear and bounded}\},$$

*with the norm as in* (5.10). *More generally, the* ***dual of a topological vector space*** *$\mathcal{X}$ is the set $\mathcal{X}^*$ of all continuous linear functionals.*

The dual $\mathcal{X}^*$ is always a complex vector space with pointwise addition and multiplication by scalars (Exercise 5.5.1). Since we have defined a norm on $\mathcal{X}^*$ when $\mathcal{X}$ is normed, the dual is itself a normed space. When $\mathcal{X}$ is a not normed, it is not entirely clear how to assign a topology to $\mathcal{X}^*$. One possibility is to consider the pointwise convergence (and we will do, and even given it a name in Section 7.2); there are, of course, other possibilities, and we will explore some of them in the future.

For the case of Hilbert spaces, we have already encountered the dual space. The Riesz Representation Theorem (4.5.4) characterizes the dual of a Hilbert space. It actually says that if $H$ is a Hilbert space, then $H^* = H$. Such an equality is hiding some information; after all, $H$ is made of vectors, whatever that means, while $H^*$ is made of linear functionals. What is missing is often call the **duality**, which is the way in which elements from one space are seen as members of the other. In this case, we identify each $\eta \in H$ with the linear functional $\hat{\eta}(\xi) = \langle \xi, \eta \rangle$.

The origin of the word "duality" probably comes from the fact that given $x \in \mathcal{X}$, $\varphi \in \mathcal{X}^*$, we can see $x$ as an element of the dual of $\mathcal{X}^*$, by $\hat{x}(\varphi) = \varphi(x)$. While this is just notation, it is a shift in point of view that is crucial in Functional Analysis: that in the expression $\varphi(x)$ we can think of the function $\varphi$ evaluating at $x$ (as is the usual notation for functions) but we can also think that we are fixing a point $x$ and "evaluating" at all $\varphi$.

At this stage we don't have enough tools to say much about duals; that will change after we study weak topologies in Chapter 7. For now we will prove the interesting fact that the dual of a normed space is a Banach space even if the original space is not Banach.

**5.5.8. Proposition.**

*For any normed vector space $\mathcal{X}$, the dual $\mathcal{X}^*$ is a Banach space.*

*Proof.* Let $\{\varphi_j\}$ be a Cauchy sequence in $\mathcal{X}^*$. For each $x \in \mathcal{X}$, the (numeric) sequence $\{\varphi_j(x)\}$ is also Cauchy, since $|\varphi_j(x)-\varphi_k(x)| \leq \|\varphi_j-\varphi_k\| \, \|x\|$; hence, convergent. Write $\varphi(x) = \lim_j \varphi_j(x)$. The linearity of $\varphi_j$ for all $j$ guarantees the linearity of $\varphi$.

Now let $\varepsilon > 0$. Then there exists $j_0$ such that $j, k > j_0$ implies $\|\varphi_j-\varphi_k\| < \varepsilon$. Also, let $m = \sup_j \|\varphi_j\|$ (since $|\|\varphi_j\| - \|\varphi_k\|| \leq \|\varphi_j - \varphi_k\|$, the numeric sequence $\{\|\varphi_j\|\}_j$ is Cauchy, so it is bounded). Then, for $k \geq j_0$

$$|\varphi(x) - \varphi_k(x)| = \lim_j |\varphi_j(x) - \varphi_k(x)| < \varepsilon \, \|x\|.$$

This shows that $\|\varphi - \varphi_k\| < \varepsilon$ if $k > j_0$. Also, for big enough $j$,

$$|\varphi(x)| \leq |\varphi(x) - \varphi_j(x)| + |\varphi_j(x)| \leq \varepsilon \, \|x\| + m \, \|x\|.$$

So $\varphi$ is bounded, i.e. $\varphi \in \mathcal{X}^*$, and it is the norm-limit of the sequence. $\square$

It is worth noting that in principle we cannot rule out the possibility that $\mathcal{X}^* = \{0\}$. While this actually happens for some TVS (cfr. Proposition 5.6.17), it is never the case for a normed space, unless $\mathcal{X} = \{0\}$; this will be a consequence of the Hahn–Banach Theorem (5.7.5).

When $\mathcal{X}$ is locally convex, we have a very similar result to Proposition 5.5.4. It is worth remembering that in an arbitrary topological space one needs to work with nets, as sequences will often be insufficient.

**5.5.9. Proposition.**

*Let $\mathcal{X}$ be a locally convex space, with the topology given by the family of seminorms $\mathcal{P}$, and $\varphi : \mathcal{X} \to \mathbb{C}$ linear. The following statements are equivalent:*

*(i)* $\varphi$ *is continuous;*

*(ii)* $\varphi$ *is continuous at* $0$*;*

*(iii)* $\varphi$ *is continuous at some point;*

*(iv)* $\ker \varphi$ *is closed;*

*(v) there exist $p_1, \ldots, p_n \in \mathcal{P}$ and $c_1, \ldots, c_n > 0$ with*

$$|\varphi(x)| \leq \sum_{j=1}^{n} c_j p_j(x), \quad x \in \mathcal{X}.$$

*When $\varphi$ is not bounded, $\ker\varphi$ is dense.*

*Proof.* (i) $\iff$ (ii) $\iff$ (iii) The proof is the same as in Proposition 5.5.4.

(i) $\implies$ (iv) We have $\ker\varphi = \varphi^{-1}(\{0\})$.

(iv) $\implies$ (ii) If $\varphi$ is not continuous at 0, there exists a net $\{x_j\}$ with $x_j \to 0$ and $|\varphi(x_j)| > \delta$ for all $j$ and a fixed $\delta > 0$. By normalizing them we can assume that $\varphi(x_j) = 1$ for all $j$, and still $x_j \to 0$. Now given $x \in \mathcal{X}$ we can form the element $x - \varphi(x)x_j \in \ker\varphi$, with $x - \varphi(x)x_j \to x$. This means that $\ker\varphi$ is dense in $\mathcal{X}$, so $\ker\varphi \subsetneq \overline{\ker\varphi}$ since $\ker\varphi$ is a proper subspace of $\mathcal{X}$. That is, $\ker\varphi$ is not closed.

(i) $\implies$ (v) Since $\varphi$ is continuous, the set $\varphi^{-1}(B_1(0))$ is open. In particular, there exists a neighbourhood $V_{p_1,\ldots,p_n;\varepsilon} \subset \varphi^{-1}(B_1(0))$, where $\varepsilon > 0$ and $p_1, \ldots, p_n \in \mathcal{P}$. This means that $x \in \mathcal{X}$ with $p_j(x) < \varepsilon$ for all $j$ implies $|\varphi(x)| < 1$. Now given $x \in \mathcal{X}$, we have that $\frac{\varepsilon x}{2\sum_{j=1}^n p_j(x)} \in V_{p_1,\ldots,p_n;\varepsilon}$. Thus

$$\left|\varphi\left(\frac{\varepsilon x}{2\sum_{j=1}^n p_j(x)}\right)\right| \leq 1,$$

which may be written as

$$|\varphi(x)| \leq \sum_{j=1}^{n} \frac{2}{\varepsilon}\, p_j(x).$$

(v) $\implies$ (ii) If $x_\ell \to 0$, then $p_j(x_\ell) \to 0$ for all $\ell$. Thus

$$|\varphi(x_\ell)| \leq \sum_{j=1}^{n} c_j p_j(x_\ell) \to 0.$$

So $\varphi$ is continuous at 0. $\square$

In particular, when $\varphi$ is an unbounded linear functional on $\mathcal{X}$, we have the unintuitive situation that the subspace $K = \ker\varphi$ is dense, and $\dim \mathcal{X}/K = 1$.

This next lemma is not about continuity, but it will be very useful nonetheless. It implies—among other things—that if the seminorms $p_j$ in Proposition 5.5.9 are of the form $p_j(x) = |\varphi_j(x)|$ for some functionals $\varphi_j$, then $\varphi$ is a linear combination of said functionals. This Lemma will be used repeatedly along the way.

**5.5.10. Lemma.**

*Let* $\varphi, \varphi_1, \ldots, \varphi_n : \mathcal{X} \to \mathbb{C}$ *be linear. If*

$$\bigcap_j \ker \varphi_j \subset \ker \varphi, \tag{5.11}$$

*then* $\varphi \in \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$.

*Proof.* Let $M \subset \mathbb{C}^n$ be given by

$$M = \{[\varphi_1(x) \ \cdots \ \varphi_n(x)]^\top : \ x \in \mathcal{X}\}.$$

Because each $\varphi_j$ is linear, $M$ is a subspace. By extending a basis of $M$ to a basis of all of $\mathbb{C}^n$, we may construct a complementary subspace $M'$ to $M$, so $\mathbb{C}^n = M \oplus M'$. Define $\psi : \mathbb{C}^n \to \mathbb{C}$ by

$$\psi([\varphi_1(x) \ \cdots \ \varphi_n(x)]^\top + m') = \varphi(x).$$

The condition (5.11) guarantees that $\psi$ is well-defined, for if $\varphi_j(y) = \varphi_j(x)$ for all $j$, then $y - x \in \bigcap_j \ker \varphi_j \subset \ker \varphi$ and so $\varphi(y) = \varphi(x)$. Since $\psi$ is linear, writing as usual $\{e_1, \ldots, e_n\}$ for the canonical basis we have

$$\varphi(x) = \psi\Big(\sum_{j=1}^n \varphi_j(x) e_j\Big) = \sum_{j=1}^n \psi(e_j)\varphi_j(x), \qquad x \in \mathcal{X}.$$

Hence $\varphi = \sum_{j=1}^n \psi(e_j)\, \varphi_j \in \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$. □

*5.5.11. Remark* . If the reader is a mere mortal like most of us, it is very likely that Lemma 5.5.10 is puzzling and does not look intuitive at all. While not claiming that the result is intuitive, let us see that it is kind of natural when $\mathcal{X}$ is a Hilbert space and the functionals are bounded. In that case, by the Riesz Representation Theorem (4.5.4) the functionals are of the form

$$\varphi_j(x) = \langle x, y_j \rangle, \qquad \varphi(x) = \langle x, y \rangle.$$

So

$$\ker \varphi_j = \{y_j\}^\perp, \qquad \ker \varphi = \{y\}^\perp.$$

Then the condition (5.11) becomes

$$\{y\}^\perp \supset \bigcap_j \{y_j\}^\perp = (\operatorname{span}\{y_1, \ldots, y_n\})^\perp.$$

Taking orthogonals,

$$\{y\} \subset \operatorname{span}\{y_1, \ldots, y_n\}.$$

This idea can be made precise in an arbitrary normed space by using *polars*. See Remark 7.3.6.

We record here a useful result, related to Lemma 5.5.10.

***5.5.12. Proposition.***

*Let $V$ be a real/complex vector space and $\varphi_1, \ldots, \varphi_n : V \to \mathbb{C}$ be linear. If $\bigcap_{j=1}^n \ker \varphi_j = \{0\}$, then $V$ is finite-dimensional. Equivalently, if $\dim V = \infty$, then $\bigcap_{j=1}^n \ker \varphi_j \neq \{0\}$ for any linear $\varphi_1, \ldots, \varphi_n : V \to \mathbb{C}$.*

*Proof.* Let $\Gamma : V \to \mathbb{C}^n$ be the linear map $\Gamma(x) = (\varphi_1(x), \ldots, \varphi_n(x))^\top$. The condition on the kernels guarantees that $\Gamma$ is injective. So $V$ is isomorphic to a subspace of $\mathbb{C}^n$, and therefore $\dim V < \infty$. □

If we look at the condition $\bigcap_{j=1}^n \ker \varphi_j = \{0\}$ in light of Lemma 5.5.10, it tells us that $\bigcap_{j=1}^n \ker \varphi_j \subset \ker \varphi$ for any linear $\varphi$, so $\varphi \in \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$. Seen this way, the condition forces the algebraic dual to be finite-dimensional, and this in turn forces $\dim V < \infty$ . We also remark that in fact, when $\dim V = \infty$, we have $\dim \bigcap_{j=1}^n \ker \varphi_j = \infty$ (Exercise 5.5.12).

**5.5.1. Exercises.**

(5.5.1) Let $\mathcal{X}$ be a normed space. Show that $\mathcal{X}^*$ is a normed vector space.

(5.5.2) Let $\mathcal{X}$ be a TVS. Show that $\mathcal{X}^*$ is a vector space.

(5.5.3) Let $\mathcal{X}$ be a finite-dimensional space. Show that $\mathcal{X}^*$ is finite-dimensional and $\dim \mathcal{X}^* = \dim \mathcal{X}$.

(5.5.4) Complete the proof of Proposition 5.5.2.

(5.5.5) Prove Proposition 5.5.4.

(5.5.6) Prove Proposition 5.5.6.

(5.5.7) Use Proposition 5.5.6 to show that (5.10) defines a norm on the space of bounded functionals on $\mathcal{X}$.

(5.5.8) Let $\mathcal{X}$ be a topological vector space, and $\varphi : \mathcal{X} \to \mathbb{C}$ linear. Suppose that there exists an open neighbourhood $V$ of 0 and $c > 0$ such that $|\varphi(v)| \leq c$ for all $v \in V$. Prove that $\varphi$ is continuous

(5.5.9) Let $\mathcal{X} = \{f \in C[0,1] : f(0) = 0\}$ with the supremum norm, and $\varphi(f) = \int_0^1 f$. Show that $\|\varphi\| = 1$ but $|\varphi(f)| < 1$ for all $f \in \mathcal{X}$ with $\|f\| = 1$.

(5.5.10) Let $\mathcal{X}$ be a normed space and $\varphi : \mathcal{X} \to \mathbb{C}$ a linear functional. Show that the following statements are equivalent:

(i) $\varphi$ is unbounded;

(ii) there exists a sequence $\{y_n\} \subset \mathcal{X}$ such that $\|y_n\| = 1$, $\varphi(y_n) > n$ for all $n$;

(iii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $x_n \to 0$, and $\varphi(x_n) = 1$ for all $n$.

(5.5.11) Let $\mathcal{X}$ be a Banach space, $\mathcal{X}_0$ a dense subspace and $\varphi : \mathcal{X}_0 \to \mathbb{C}$ a bounded linear functional. Show that $\varphi$ admits a unique extension $\tilde{\varphi} \in \mathcal{X}^*$.

(5.5.12) Let $V$ be an infinite-dimensional real/complex vector space. Consider linear maps $\varphi_1, \ldots, \varphi_n : V \to \mathbb{C}$. Improve on Proposition 5.5.12 by showing that

$$\dim \bigcap_{j=1}^{n} \ker \varphi_j = \infty.$$

## 5.6. Examples of Duals

In this section we stop the development of the theory momentarily, to discuss some examples. The reader can safely skip to Section 5.7 and come back here any time, though the examples will likely help in getting some intuition.

We will look at some well-known cases where it is possible to identify the dual of the Banach space explicitly. Doing so will also give us an opportunity to discuss some of the best-known Banach spaces. A basic knowledge of measure theory will be assumed (cfr. Chapter 2; most details about $L^p$ spaces are developed in Section 2.8).

Let us now move into considering the examples. We list them here, for easy access:



**5.6.1.** $\ell^p(\mathbb{N})$**,** $1 \leq p < \infty$**.** In this case our locally convex space is the (Banach) space

$$\ell^p(\mathbb{N}) = \Big\{ f : \mathbb{N} \to \mathbb{C} : \sum_{k=1}^{\infty} |f(k)|^p < \infty \Big\},$$

with the norm $\|f\|_p = \left(\sum_{k=1}^\infty |f(k)|^p\right)^{1/p}$. These are just particular cases of the $L^p(X)$ spaces, with the counting measure. We provide details for this case though, for the benefit of the reader who is not proficient in measure theory—such reader should consider addressing this shortage rather sooner than later!

We recall that in $\ell^p(\mathbb{N})$ we have the **canonical basis** $\{e_k\}$ of $\ell^p(\mathbb{N})$; for each $k$, $e_k$ is the characteristic of $k$, i.e. $e_k(j) = \delta_{k,j}$. It is a Schauder basis (Exercise 5.1.6). We also note that the results below also apply to $\ell^p(S)$ for any nonempty set $S$; we focus on $S = \mathbb{N}$ for simplicity, and because it is by far the most common case.

This next result is particular case of Proposition 5.6.8.

***5.6.1. Proposition.***

*Let $q \in [1, \infty]$ and $p \in \mathbb{R}$ with $\frac{1}{p} + \frac{1}{q} = 1$, $g : \mathbb{N} \to \mathbb{C}$. Then*

$$\|g\|_q = \max\left\{ \left|\sum_{k=1}^\infty g(k)f(k)\right| \;:\; f \in \ell^p(\mathbb{N}),\ \|f\|_p = 1 \right\}. \tag{5.12}$$

*The equality holds even when one (and thus the other) of the sides is infinite.*

*Proof.* If $g = 0$, the equality holds trivially, so we take $g \neq 0$. We consider first that $q < \infty$. The proof works even when $q = 1$, $p = \infty$; in such case we use the convention $q/p = 0$. Assume first that $\sum_k |g(k)|^q < \infty$. Then, for any $f \in \ell^p(\mathbb{N})$ with $\|f\|_p = 1$, we apply Hölder's inequality to get

$$\sum_{k=1}^\infty |g(k)f(k)| \leq \|g\|_q.$$

This shows that the right-hand-side in (5.12) is less than or equal the left hand side. Now, for each $k$, let $g(k) = |g(k)|e^{i\theta_k}$ be the polar form. Let $f_0(k) = e^{-i\theta_k}|g(k)|^{q-1}$. Then, if $p < \infty$,

$$\sum_{k=1}^\infty |f_0(k)|^p = \sum_{k=1}^\infty |g(k)|^{p(q-1)} = \sum_{k=1}^\infty |g(k)|^q < \infty.$$

And if $p = \infty$, then $q = 1$ and $|f_0(k)| = 1$ for all $k$. In both cases, $f_0 \in \ell^p(\mathbb{N})$. Let $f_1 = f_0/\|g\|_q^{q/p}$. Then $\|f_1\|_p = 1$. We have

$$\begin{aligned}\sum_{k=1}^\infty g(k)f_1(k) &= \frac{1}{\|g\|_q^{q/p}} \sum_{k=1}^\infty g(k)e^{-i\theta_k}|g(k)|^{q-1} = \frac{1}{\|g\|_q^{q/p}} \sum_{k=1}^\infty |g(k)|^q \\ &= \|g\|_q^{q-q/p} = \|g\|_q,\end{aligned}$$

proving the reverse inequality.

Now suppose that $\sum_{k=1}^{\infty} |g(k)|^q = \infty$. For each $n \in \mathbb{N}$, the function $f_n$, given by

$$f_n(k) = \frac{e^{-i\theta_k} |g(k)|^{q-1}}{\left(\sum_{k=1}^{n} |g(k)|^q\right)^{1/p}} 1_{[1,n]},$$

is in $\ell^p(\mathbb{N})$ and $\|f_n\|_p = 1$. Also,

$$\sum_{k=1}^{n} g(k) f_n(k) = \left(\sum_{k=1}^{n} |g(k)|^q\right)^{1-1/p} = \left(\sum_{k=1}^{n} |g(k)|^q\right)^{1/q}$$

is arbitrarily large as $n$ grows, so the supremum on the right-hand-side in (5.12) is infinite too.

When $q = \infty$, fix $\varepsilon > 0$. Then there exists an index $j$ with $|g(j)| + \varepsilon > \|g\|_\infty$. Let $g(j) = |g(j)| e^{i\theta}$. Then, with $f = e^{-i\theta} e_j$,

$$\sum_{k=1}^{\infty} g(k) f(k) = g(j) e^{-i\theta} = |g(j)| > \|g\|_\infty - \varepsilon.$$

As $\varepsilon$ was arbitrary, this proves the equality. □

*5.6.2. Remark.* The equality (5.12) can be also written as

$$\|g\|_q = \sup\left\{ \sum_{k=1}^{\infty} |g(k) f(k)| \ : \ f \in \ell^p(\mathbb{N}),\ \|f\|_p = 1 \right\} \tag{5.13}$$

(Exercise 5.6.1).

Now we will use Proposition 5.6.1 to determine the dual of $\ell^p(\mathbb{N})$. Namely,

***5.6.3. Proposition.***

*Let $1 \le p < \infty$. Then the dual of $\ell^p(\mathbb{N})$ is isometrically isomorphic to $\ell^q(\mathbb{N})$, where $q$ is such that $\frac{1}{p} + \frac{1}{q} = 1$. The isomorphism is the map $\gamma : \ell^q(\mathbb{N}) \to \ell^p(\mathbb{N})^*$ given by $\gamma(g)(f) = \langle f, g\rangle$.*

*Proof.* Consider first the case where $p > 1$. Let $\varphi \in \ell^p(\mathbb{N})^*$. Then $\varphi$ is bounded, so there exists a constant $c$ with $|\varphi(f)| \le c\|f\|_p$ for all $f \in \ell^p(\mathbb{N})$. Let $n \in \mathbb{N}$. Then

$$\left|\sum_{k=1}^{n} \varphi(e_k) f(k)\right| = \left|\varphi\left(\sum_{k=1}^{n} f(k) e_k\right)\right| \le c \left(\sum_{k=1}^{n} |f(k)|^p\right)^{1/p} \le c\,\|f\|_p.$$

As $n$ and $f$ were arbitrary, Proposition 5.6.1 implies that the function $g_\varphi : k \mapsto \varphi(e_k)$ is in $\ell^q(\mathbb{N})$. Noting that $f = \lim_n \sum_{k=1}^{n} f(k) e_k$ in $\ell^p(\mathbb{N})$ since the canonical basis is Schauder and using that $\varphi$ is bounded for the first equality, we have

$$\varphi(f) = \lim_{n\to\infty} \varphi\left(\sum_{k=1}^{n} f(k)e_k\right) = \lim_{n\to\infty} \sum_{k=1}^{n} f(k)\varphi(e_k)$$
$$= \sum_{k=1}^{\infty} f(k)\, g_\varphi(k) = \langle f, g_\varphi\rangle.$$

So every $\varphi \in \ell^p(\mathbb{N})^*$ is determined by a function in $\ell^q(\mathbb{N})$, in the sense that $\varphi = \langle\,\cdot\,, g_\varphi\rangle$.

Conversely, if $g \in \ell^q(\mathbb{N})$, the map

$$f \mapsto \sum_{k=1}^{\infty} f(k)g(k) \tag{5.14}$$

defines a bounded functional with norm $\|g\|_q$ (Exercise 5.6.7). So the map $\varphi \longmapsto \{\varphi(e_k)\}_k$ is an isometric isomorphism of $\ell^p(\mathbb{N})^*$ onto $\ell^q(\mathbb{N})$. The inverse of this map is $\gamma$ as in the statement of the Proposition.

We leave the case $p = 1$ as an exercise (Exercise 5.6.6). □

At this stage we can use Proposition 5.5.8 to conclude that $\ell^p(\mathbb{N})$ is a Banach space. It can also be seen as a direct consequence of Theorem 2.8.12.

*5.6.4. Remark*. We will sometimes write the assignment from equation (5.14) as $f \mapsto \langle f, g\rangle$. Note that in the case $p = q = 2$ (i.e. Hilbert space), we get the actual inner product. In this notation Hölder's inequality is written as $|\langle f, g\rangle| \leq \|f\|_p \, \|g\|_q$.

Here is a result, characterizing the dual of a direct sum, that will be needed later.

**5.6.5. Proposition.**

*Let $\mathcal{X}_1, \mathcal{X}_2, \dots$ be normed spaces, $p \in [1, \infty)$. Then there exists a canonical isometric isomorphism $\Gamma : \left(\bigoplus_n \mathcal{X}_n\right)_{\ell^p}^* \to \left(\bigoplus_n \mathcal{X}_n^*\right)_{\ell^q}$. When the direct sum is finite, the isomorphism also occurs for $p = \infty$.*

*Proof.* Given $\varphi \in \left(\bigoplus_n \mathcal{X}_n\right)_{\ell^p}^*$ we write $\varphi_n$ to denote the restriction to the $n^{\text{th}}$ component $\mathcal{X}_n$. Since the natural inclusion $\mathcal{X}_m \hookrightarrow \left(\bigoplus_n \mathcal{X}_n\right)_{\ell^p}$ is isometric, we get that $\varphi_m \in \mathcal{X}_m^*$.

We want to define $\Gamma(\varphi) = \bigoplus_n \varphi_n$. The first issue is to check that $\Gamma(\varphi)$ is indeed in $\Big(\bigoplus_n \mathcal{X}_n^*\Big)_{\ell^q}$. Fix $0 < c < 1$, and let $z_n \in \mathcal{X}_n$ with $\|z_n\| = 1$ and $\varphi_n(z_n) \geq c\,\|\varphi_n\|$ (we can always get the left-hand-side to be nonnegative by multiplying $z_n$ by a suitable number in $\mathbb{T}$). Define, for each $m \in \mathbb{N}$ (and when $q < \infty$, otherwise we tweak the definition slightly),

$$w_m = \bigoplus_{n=1}^{m} \frac{\|\varphi_n\|^{q/p}}{\left(\sum_{k=1}^m \|\varphi_n\|^q\right)^{1/p}}\, z_n \in \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}$$

By construction, $\|w_m\| = 1$ and, since $\frac{1}{q} = 1 - \frac{1}{p}$ and $q = \frac{q}{p} + 1$,

$$\begin{aligned}\Big(\sum_{n=1}^m \|\varphi_n\|^q\Big)^{1/q} &= \sum_{n=1}^m \frac{\|\varphi_n\|^q}{\left(\sum_{n=1}^m \|\varphi_n\|^q\right)^{1/p}} = \sum_{n=1}^m \frac{\|\varphi_n\|^{q/p}}{\left(\sum_{n=1}^m \|\varphi_n\|^q\right)^{1/p}}\, \|\varphi_n\| \\ &\leq \sum_{n=1}^m \frac{\|\varphi_n\|^{q/p}}{\left(\sum_{n=1}^m \|\varphi_n\|^q\right)^{1/p}}\, c^{-1}\varphi_n(z_n) = c^{-1}\,\varphi(w_m) \\ &\leq c^{-1}\|\varphi\|.\end{aligned}$$

Letting $c \to 1$ and $m \to \infty$ shows that $\Gamma$ is well-defined and that $\|\Gamma\| \leq 1$.

Next we define $\Gamma' : \Big(\bigoplus_n \mathcal{X}_n^*\Big)_{\ell^q} \to \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}^*$ by

$$\Gamma'\Big(\bigoplus_n \varphi_n\Big)\,\big(\bigoplus_n x_n\big) = \sum_n \varphi_n(x_n).$$

We have

$$\begin{aligned}\Big|\Gamma'\Big(\bigoplus_n \varphi_n\Big)\,\big(\bigoplus_n x_n\big)\Big| &= \Big|\sum_n \varphi_n(x_n)\Big| \leq \sum_n \|\varphi_n\|\,\|x_n\| \\ &\leq \Big(\sum_n \|\varphi_n\|^q\Big)^{1/q}\Big(\sum_n \|x_n\|^p\Big)^{1/p} \\ &= \Big\|\bigoplus_n \varphi_n\Big\|\,\Big\|\bigoplus_n x_n\Big\|.\end{aligned}$$

This shows that $\Gamma'$ maps into the right space, and also that $\|\Gamma'\| \leq 1$. Since $\Gamma' \circ \Gamma = \mathrm{id}_{(\bigoplus_n \mathcal{X}_n)^*_{\ell^p}}$ and $\Gamma \circ \Gamma' = \mathrm{id}_{(\bigoplus_n \mathcal{X}_n^*)_{\ell^q}}$, we get that $\Gamma$ and its inverse are contractions, so $\Gamma$ is isometric. $\square$

**5.6.2.** $c_0$**.** This is the space of sequences converging to zero. In this case, since the canonical basis is a Schauder basis for $c_0$ (Exercise 5.1.9), the proof of Proposition 5.6.3 does not break down if we use $c_0$ instead of $\ell^\infty(\mathbb{N})$, and so we can show that ${c_0}^* = \ell^1(\mathbb{N})$.

**5.6.3. $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$.** There is a reason why we did not prove that the dual of $\ell^\infty(\mathbb{N})$ is $\ell^1(\mathbb{N})$: it is not. The argument that proves that the dual of $\ell^p(\mathbb{N})$ is $\ell^q(\mathbb{N})$ $(1 \leq p < \infty)$ breaks down when $p = \infty$. Indeed, a key piece of the argument is the fact that the canonical basis is a Schauder basis for $\ell^p(\mathbb{N})$ (in the assertion "$f = \lim_n \sum_{k=1}^n f(k) e_k$"); this is not true in $\ell^\infty(\mathbb{N})$. So all we can prove is with the above techniques is

**5.6.6. Proposition.**

*The dual of $\ell^\infty(\mathbb{N})$ contains $\ell^1(\mathbb{N})$.*

The inclusion $\ell^1(\mathbb{N}) \subset \ell^\infty(\mathbb{N})^*$ can be shown to be proper. There are several ways to see this, all requiring a bit more machinery than currently available. A more or less straightforward argument is laid out in Exercise 5.7.4. A second way is to notice that $\ell^1(\mathbb{N})$ is separable while $\ell^\infty(\mathbb{N})$ is not; then Corollary 5.7.11 shows that $\ell^\infty(\mathbb{N})^*$ is not separable, so it cannot be isomorphic to $\ell^1(\mathbb{N})$. And yet another way is to let $\omega$ be any free ultrafilter (see Chapter E for information on ultrafilters) on $\mathbb{N}$. Then we can define a bounded linear functional $\varphi : \ell^\infty(\mathbb{N}) \to \mathbb{C}$ by

$$\varphi(f) = \lim_{n \to \omega} f(n)$$

This functional cannot come from any $g \in \ell^1(\mathbb{N})$ for the following reason: it is zero on any $f \in c_{00}$, and no nonzero $g \in \ell^1(\mathbb{N})$ can have the property that $\langle f, g \rangle = 0$ for all $f \in c_{00}$.

The dual of $\ell^\infty(\mathbb{N})$ can be explicitly characterized, even if said characterization is not particularly helpful in the sense that the previous explicit characterizations of duals are. The proof below works in way more generality with basically the same idea; we refer the interested reader to [DS88, Theorem IV.8.16].

We denote by $BA(\mathbb{N})$ the set

$$BA(\mathbb{N}) = \{b : \mathcal{P}(\mathbb{N}) \to [0, \infty], \quad \text{bounded and additive}\},$$

with the total variation as the norm (cfr. Definition 2.10.1). The elements of $BA(\mathbb{N})$ are not necessarily $\sigma$-additive.

**5.6.7. Proposition.**

*The dual of $\ell^\infty(\mathbb{N})$ is canonically isometrically isomorphic to the space $BA(\mathbb{N})$.*

*Proof.* We define $\Gamma : BA(\mathbb{N}) \to \ell^\infty(\mathbb{N})^*$ by $\Gamma(b)(x) = \int_{\mathbb{N}} x \, db$. Though $b$ is not $\sigma$-additive, it behaves as a measure in every other respect and we can define the Lebesgue integral.

Given disjoint finite $E_1, \ldots, E_n \subset \mathbb{N}$ such that $\sum_j |b(E_j)| \geq \|b\| - \varepsilon$, choose $\alpha_1, \ldots, \alpha_n \in \mathbb{T}$ with $\alpha_j\, b(E_j) = |b(E_j)|$ and define $x = \sum_j \alpha_j\, 1_{E_j}$. Then $x \in \ell^\infty(\mathbb{N})$, $\|x\| = 1$, and

$$\|\Gamma(b)\| \geq |\Gamma(b)(x)| = \left| \int_{\mathbb{N}} x\, db \right| = \Big| \sum_j \alpha_j\, b(E_j) \Big| = \sum_j |b(E_j)| \geq \|b\| - \varepsilon.$$

This works for any $\varepsilon > 0$, so $\|b\| \leq \|\Gamma(b)\|$. It is tempting to use familiar results from Measure Theory to show that $|\Gamma(b)(x)| \leq \|x\|\,\|b\|$ for all $x$. The problem is that our measure is not necessarily $\sigma$-additive, and so we cannot use the familiar results; concretely, the $\sigma$-additivity of the measure is used to show continuity of the measure which then is crucial in the Monotone Convergence Theorem. So those are out of the question. But, more elementarily, for any simple $s = \sum_j x_j\, 1_{E_j}$ that approximates $x$ (so we can choose $x_j$ with $|x_j| \leq \|x\|_\infty$),

$$\left| \int_{\mathbb{N}} s\, db \right| = \Big| \sum_j x_j\, b(E_j)| \leq \|x\|_\infty \sum_j |b(E_j)| \leq \|x\|_\infty\, \|b\|.$$

And as this holds for any such $s$, by the definition of the Lebesgue integral we get $|\Gamma(b)(x)| \leq \|x\|_\infty\, \|b\|$, which in turn gives us $\|\Gamma(b)\| \leq \|b\|$. So $\Gamma$ is isometric, and all it remains is to show that it is surjective. Given $\varphi \in \ell^\infty(\mathbb{N})^*$, define $b \in BA(\mathbb{N})$ by $b(E) = \varphi(1_E)$. If $x = \sum_{j=1}^m \beta_j\, 1_{E_j}$, then by the finite additivity of $b$ we have

$$\Gamma(b)(x) = \sum_{j=1}^m \beta_j\, b(E_j) = \sum_{j=1}^m \beta_j\, \varphi(1_{E_j}) = \varphi\Big( \sum_{j=1}^m \beta_j\, 1_{E_j} \Big) = \varphi(x).$$

As simple functions are uniformly dense in $\ell^\infty(\mathbb{N})$ (Theorem 2.4.13), given $x \in \ell^\infty(\mathbb{N})$ we get a sequence $\{x_n\}$ of simple functions in $\ell^\infty(\mathbb{N})$ such that $\|x_n - x\|_\infty \to 0$. Then

$$\left| \int_{\mathbb{N}} |x_n - x|\, db \right| \leq \|x_n - x\|_\infty\, \|b\| \to 0,$$

which implies that $\Gamma(b)(x) = \lim_n \Gamma(b)(x_n) = \lim_n \varphi(x_n) = \varphi(x)$. Hence $\Gamma$ is surjective. $\square$

For $1 < p < \infty$, the second dual of $\ell^p(\mathbb{N})$ is canonically isometrically isomorphic to $\ell^p(\mathbb{N})$, which implies that $\ell^p(\mathbb{N})$ is **reflexive**. This is not true for $\ell^1(\mathbb{N})$ nor $\ell^\infty(\mathbb{N})$.

Let us also mention that, although it is in some sense a very particular example of a Banach space, the space $\ell^1(\mathbb{N})$ is in another sense universal among separable Banach spaces (Theorem 6.1.10).

**5.6.4. Hilbert spaces.** By the Riesz Representation Theorem 4.5.4, there is an isometric isomorphism $\mathcal{H} \to \mathcal{H}^*$, with the duality given by $\varphi_\xi(\eta) = \langle \eta, \xi \rangle$

(for the case of countable dimension, this is the situation where $p = q = 2$ in subsection 5.6.1 above). Properly, this natural isomorphism is conjugate-linear rather than linear, but this distinction makes no difference in most situations.

**5.6.5. $L^p(X, \Omega, \mu)$, $1 \le p < \infty$.** Let $(X, \Omega, \mu)$ be a measure space. Recall that a measurable function $f$ is **finitely supported** if $\mu\{|f| > 0\} < \infty$. Depending on the measure space, $L^p(X)$ can be fairly pathological. For instance, if $\mu(E) = \infty$ for all $E \subset X$, then $L^p(X, \mu) = \{0\}$. We need to avoid those cases, hence the $\sigma$-finiteness requirement.

The next few results have already been proven above in the discrete case. We can prove them in the general case with little change.

***5.6.8. Proposition.***

*Let $(X, \Omega, \mu)$ be a measure space, $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, $g : X \to \mathbb{C}$ measurable. Then*

$$\|g\|_q = \sup \left\{ \left| \int_X gf \, d\mu \right| : \; f \in L^p(X), \; \|f\|_p = 1 \right\}. \tag{5.15}$$

*The equality holds even when one (and thus the other) of the sides is infinite, as long as $(X, \Omega, \mu)$ is $\sigma$-finite. The equality for $p = 1$, $q = \infty$ also requires $\sigma$-finiteness, even when both sides of the equality are finite.*

*If $\|g\|_q < \infty$, the equality (5.15) still holds if we only allow simple functions on the right-hand-side. When $\|g\|_q = \infty$ and $\mu$ is $\sigma$-finite, the equality also holds with simple functions on the right-hand-side.*

*Proof.* We do first the case $1 < p < \infty$. The right-hand-side in (5.15) is less than or equal the left hand side by Hölder's inequality. For the reverse inequality, assume first that $\int_X |g|^q \, d\mu < \infty$. Write $g(x) = |g(x)| e^{i\theta(x)}$ the polar form; as $\theta(x) = \arctan \operatorname{Im} g(x) / \operatorname{Re} g(x)$, we see that the function $\theta$ is measurable. Let $f_0 = e^{-i\theta} |g|^{q-1}$. Then

$$\int_X |f_0|^p \, d\mu = \int_X |g|^{p(q-1)} \, d\mu = \int_X |g|^q \, d\mu < \infty.$$

So $f_0 \in L^p(X)$. Let $f_1 = f_0 / \|g\|_q^{q/p}$. Then $\|f_1\|_p = 1$. We have

$$\begin{aligned} \int_X g f_1 \, d\mu &= \frac{1}{\|g\|_q^{q/p}} \int_X g \, e^{-i\theta} \, |g|^{q-1} \, d\mu = \frac{1}{\|g\|_q^{q/p}} \int_X |g|^q \\ &= \|g\|_q^{q - q/p} = \|g\|_q. \end{aligned}$$

Now suppose that $\int_X |g|^q \, d\mu = \infty$ and that $\mu$ is $\sigma$-finite. If $\mu(\{|g| = \infty\}) > 0$, we can take $f$ to be the simple function $1_{\{|g|=\infty\}}$. Otherwise, $|g| < \infty$ a.e. By the $\sigma$-finiteness there exists $\{X_n'\}$ increasing, measurable, with $\bigcup_n X_n' = X$.

Let $X_n = X'_n \cap \{|g| \leq n\}$. Then $X = \bigcup_n X_n$ is an increasing union of finite-measure sets, and $\int_X |g|^q \, d\mu = \lim_n \int_{X_n} |g|^q \, d\mu$, with $\int_{X_n} |g|^q < \infty$ for all $n$. Define

$$f_n = \frac{e^{-i\theta}|g|^{q-1}}{\left(\int_{X_n} |g|^q \, d\mu\right)^{1/p}} 1_{X_n}.$$

Then $f_n \in L^p(X)$ and $\|f_n\|_p = 1$. Also,

$$\int_X g \, f_n \, d\mu = \left(\int_{X_n} |g|^q \, d\mu\right)^{1-1/p} = \left(\int_{X_n} |g|^q \, d\mu\right)^{1/q}$$

is arbitrarily large as $n$ grows, so the supremum on the right-hand-side in (5.15) is infinite too. If we choose $h_n$ to be simple and such that

$$\|f_n - h_n\|_p < \left(\int_{X_n} |g|^q \, d\mu\right)^{-1/q},$$

then

$$\begin{aligned}\left|\int_X g \, h_n \, d\mu\right| &= \left|\int_{X_n} g \, h_n \, d\mu\right| \\ &\geq \int_{X_n} g \, f_n \, d\mu - \left(\int_{X_n} |g|^q \, d\mu\right)^{1/q} \|f_n - h_n\|_p \\ &\geq \int_{X_n} g \, f_n \, d\mu - 1,\end{aligned}$$

still growing arbitrarily with $n$.

When $p = 1$ and $\mu$ is $\sigma$-finite, suppose first that $\|g\|_\infty < \infty$. By Hölder, if $\|f\|_1 = 1$ we have $\int_X |fg| \, d\mu \leq \|g\|_\infty$. Conversely, let $\varepsilon > 0$. Then there exists $E \subset X$, measurable, with $0 < \mu(E) < \infty$ and $|g| \geq \|g\|_\infty - \varepsilon$ on $E$. Let $f = \frac{1}{\mu(E)} 1_E$. Then $\|f\|_1 = 1$, and

$$\int_X |fg| \, d\mu = \frac{1}{\mu(E)} \int_E |g| \, d\mu \geq (\|g\|_\infty - \varepsilon).$$

As we can do this for any $\varepsilon > 0$, we obtain (5.15). As the $f$ we used is simple, the equality holds when we restrict to simple functions.

If $\|g\|_\infty = \infty$, then for each $n \in \mathbb{N}$ there exists $E_n \subset X$, measurable, with $0 < \mu(E_n) < \infty$ (the right bound here is where we use $\sigma$-finite) and $|g| \geq n$ on $E_n$. Let $f_n = \frac{1}{\mu(E_n)} 1_{E_n}$. Then $\|f_n\|_1 = 1$ for all $n$, and

$$\int_X |f_n g| \, d\mu = \frac{1}{\mu(E_n)} \int_{E_n} |g| \, d\mu \geq n \, \frac{1}{\mu(E_n)} \int_{E_n} 1 \, d\mu = n,$$

so the right-hand-side in (5.15) is infinite.

Finally, when $p = \infty$, if $\|f\|_\infty = 1$ we have

$$\int_X |gf| \, d\mu \leq \int_X |g| \, d\mu = \|g\|_1,$$

with equality for $f = 1$. If $\|g\|_1 = \infty$, the right-hand-side in (5.22) is also infinite by taking $f = 1$.

We finish by addressing that we can restrict to simple functions in the right-hand-side of (5.15). If $\|g\|_q < \infty$ and $f \in L^p(X)$ with $\|f\|_p = 1$ and $\int_X fg = \|g\|_q$, then by Proposition 2.8.17 for a fixed $\varepsilon > 0$ there exists $f_0 \in L^p(X)$, simple, with $\|f - f_0\|_p < \varepsilon$. Let $f_1 = f_0/\|f_0\|_p$, still simple. As $|\|f\|_p - \|f_0\|_p| \leq \|f - f_0\|_p < \varepsilon$ and $\|f\|_p = 1$, we have that $1 - \varepsilon \leq \|f_0\|_p \leq 1 + \varepsilon$. So

$$\begin{aligned}\|f - f_1\|_p = \frac{\big\|f\,\|f_0\|_p - f_0\big\|_p}{\|f_0\|_p} &\leq \frac{\|f - f_0\|_p\,\|f_0\|_p + \big|\,1 - \|f_0\|_p\,\big|\,\|f_0\|_p}{\|f_0\|_p} \\ &= \|f - f_0\|_p + |\,1 - \|f_0\|_p| < \varepsilon + \varepsilon = 2\varepsilon.\end{aligned}$$

Now

$$\begin{aligned}\left|\int_X gf_1\right| \geq \left|\int_X gf\right| - \left|\int_X g(f_1 - f)\right| &= \|g\|_q - \left|\int_X g(f_1 - f)\right| \\ &\geq \|g\|_q - \|g\|_q\,\|f_1 - f\|_p \geq (1 - 2\varepsilon)\|g\|_q.\end{aligned}$$

As $\varepsilon$ was arbitrary, this shows that (5.15) holds when allowing only simple $f \in L^p(X)$. $\square$

*5.6.9. Remark*. The equality (5.15) can also be written as

$$\|g\|_q = \sup\left\{\int_X |gf|\,d\mu : \ f \in L^p(X),\ \|f\|_p = 1\right\}. \tag{5.16}$$

This follows directly from the fact that $|gf| = ghf$ for an appropriate measurable $h$ with $|h| = 1$, and $\|hf\|_p = \|f\|_p$.

*5.6.10. Remark*. To see what can go wrong when $X$ is not $\sigma$-finite in Proposition 5.6.8, consider as in Exercise 5.6.10 the space $X = \{0, 1\}$ with the measure $\mu(\{0\}) = 1$, $\mu(\{1\}) = \infty$. Then $L^p(\mu) = \{h : \ h(1) = 0\}$ for all $p < \infty$. If we take $g : \{0, 1\} \to \mathbb{C}$ given by $g(0) = 0$, $g(1) = 1$, then $\|g\|_\infty = 1$, while $\int_X gf\,d\mu = f(0)g(0) = 0$ for all $f \in L^1(\mu)$; we also have that $\|g\|_q = \infty$ for any $q < \infty$, and $\int_X gf\,d\mu = 0$ for all $f \in L^p(\mu)$. So (5.15) fails in general when $q = \infty$, $p = 1$, and in the infinite case for arbitrary $q$.

The argument used in Proposition 5.6.1 to determine the dual of $\ell^p(\mathbb{N})$ doesn't extend directly to the general case. The spirit is the same, though.

**5.6.11. Theorem.**

*Let $(X, \Omega, \mu)$ be a measure space and $1 < p < \infty$. Then the dual of $L^p(X)$ is isometrically isomorphic to $L^q(X)$, where $q$ is such that $\frac{1}{p} + \frac{1}{q} = 1$. The isomorphism is given by $\Gamma : L^q(X) \to L^p(X)^*$,*

$$\Gamma(g) f = \int_X f\, g\, d\mu. \tag{5.17}$$

*If $\mu$ is $\sigma$-finite, the duality also holds for $p = 1$.*

*Proof.* A short argument shows that $\Gamma$ as in (5.17) is linear and bounded (Exercise 5.6.16). So the rest of the proof is devoted to showing that $\Gamma$ is surjective and isometric.

Assume first that $p \in [1, \infty)$ and $\mu(X) < \infty$. Let $\phi \in L^p(X)^*$. Define, for $E \subset X$ measurable, $\mu_\phi(E) = \phi(1_E)$. This is well-defined since $1_E \in L^p(X)$ (from $\mu(X) < \infty$). If $\{E_n\}$ is a pairwise disjoint sequence of measurable subsets of $X$, note that $\sum_n \mu(E_n) = \mu(\bigcup_n E_n) < \infty$. Then

$$\begin{aligned}\mu_\phi\Big(\bigcup_{n>m} E_n\Big) &\le \|\phi\|\, \|1_{\bigcup_{n>m} E_n}\|_p = \|\phi\| \Big(\sum_{n>m} \mu(E_n)\Big)^{1/p} \\ &\le |\phi\| \Big(\sum_{n>m} \mu(E_n)\Big)^{1/p} \xrightarrow[m\to\infty]{} 0.\end{aligned}$$

So

$$\mu_\phi\Big(\bigcup_n E_n\Big) = \phi(1_{\bigcup_n E_n}) = \lim_{m\to\infty} \sum_{n=1}^m \mu_\phi(E_n) = \sum_{n=1}^\infty \mu_\phi(E_n).$$

Thus $\mu_\phi$ is a measure, absolutely continuous with respect to $\mu$. Let $g \in L^1(X)$ be its Radon–Nikodym derivative. We have, for any simple function $f = \sum_j \alpha_j\, 1_{E_j}$,

$$\phi(f) = \sum_j \alpha_j \mu_\phi(E_j) = \int_X f\, d\mu_\phi = \int_X f\, g\, d\mu. \tag{5.18}$$

By Proposition 5.6.8,

$$\|g\|_q = \sup\{|\phi(f)| : \ f \in L^p(X) \ \text{ simple}, \ \|f\|_p = 1\} \le \|\phi\|. \tag{5.19}$$

So $g \in L^q(X)$. Now if $f \in L^p(X)$ and $\{f_k\} \subset L^p(X)$ are simple with $\|f - f_k\|_p \to 0$ (via Proposition 2.8.17), we have $\phi(f) = \lim_k \phi(f_k)$ and, by Hölder,

$$\left| \int_X fg\, d\mu - \int_X f_k g\, d\mu \right| \le \|f - f_k\|_p\, \|g\|_q \to 0.$$

Thus (5.18) holds for all $f \in L^p(X)$. So $\|g\|_q = \sup\{|\varphi(f)| : \ \|f\|_p = 1\} = \|\varphi\|$. And Proposition 5.6.8 also gives us that $g$ is unique for a given $\phi$.

When $\mu(X) = \infty$ and $\mu$ is $\sigma$-finite, let $\{X_n\}$ be measurable, increasing, finite measure, and $X = \bigcup_n X_n$. Let $\phi \in L^p(X)^*$. Embedding $L^p(X_n)$ naturally into $L^p(X)$, we can consider $\phi$ restricted to $L^p(X_n)$ and apply the first part of the proof to obtain a unique $g_n \in L^q(X_n)$ with $\|g_n\|_q \leq \|\phi\|$, and

$$\phi(f) = \int_{X_n} f\, g_n \, d\mu$$

for all $f \in L^p(X)$ supported on $X_n$. By the uniqueness in the first part, we get that $g_{n+1}|_{X_n} = g_n$. So we may define $g$ by $g|_{X_n} = g$. Since $X_n \subset X_{n+1}$ for all $n$, by Monotone Convergence

$$\int_X |g|^q \, d\mu = \lim_n \int_{X_n} |g|^q \, d\mu = \lim_n \int_{X_n} |g_n|^q \, d\mu \leq \|\phi\|,$$

so $g \in L^q(X)$. With the same idea we show that $\phi(f) = \int_X fg \, d\mu$ for all $f \in L^p(X)$, and then Proposition 5.6.8 gives us that $\|g\|_q = \|\phi\|$.

When $X$ is not necessarily $\sigma$-finite and $p > 1$, the key idea is that for nonzero $f \in L^p(X)$, we need $\mu(\{|f| > c\}) < \infty$ for any $c > 0$. So the set $Z_f = \{f \neq 0\}$ is $\sigma$-finite; the problem we have is that $Z_f$ depends on $f$, so we need to somehow amalgamate all the $Z_f$, $f \in L^p(X)$. Let $E \subset X$ be $\sigma$-finite. By the previous part of the proof, there exists $g_E \in L^q(X)$ with $\phi(f) = \int_E fg_E \, d\mu$ for all $f \in L^p(E)$ and $\|g_E\|_q \leq \|\phi\|$. Also, if $F \supset E$ is another $\sigma$-finite set, we get that $g_F|_E = g_E$. Let

$$s = \sup\{\|g_E\|_q : \ E \subset X, \ \sigma\text{-finite}\} \leq \|\phi\|.$$

This allows us to choose a sequence $\{E_n\}$ of $\sigma$-finite sets $E_n$ such that $s = \lim_n \|g_{E_n}\|_q$. Put $E = \bigcup_n E_n$; this set is $\sigma$-finite, and $\|g_E\|_q \geq \|g_{E_n}\|_q$ for all $n$, so $\|g_E\|_q = s$. If $F$ is $\sigma$-finite and $F \supset E$, note that $F \setminus E$ is a $\sigma$-finite subset of $F$, so $g_{F\setminus E} = g_F|_{F\setminus E}$. Then

$$\begin{aligned}\int_X |g_E|^q \, d\mu + \int_X |g_{F\setminus E}|^q \, d\mu &= \int_E |g_F|^q \, d\mu + \int_{F\setminus E} |g_F|^2 \, d\mu \\ &= \int_X |g_F|^q \, d\mu \leq s^q = \int_X |g_E|^q \, d\mu,\end{aligned}$$

so $g_{F\setminus E} = 0$ a.e.; thus $g_F = g_E + g_{F\setminus E} = g_E$. Now, given $f \in L^p(X)$, we may consider the $\sigma$-finite set $F = E \cup \{f \neq 0\}$; this set $F$ depends on $f$. Then by the $\sigma$-finite proof,

$$\phi(f) = \int_F f\, g_F \, d\mu = \int_F f\, g_E \, d\mu = \int_X f\, g_E \, d\mu.$$

Thus $g = g_E \in L^q(X)$ is the desired function that gives us $\phi(f) = \int_X fg \, d\mu$ for all $f \in L^p(X)$. And $\|g\|_q = \|\phi\|$ as before, since $p > 1$ and Proposition 5.6.8 still applies. $\square$

*5.6.12. Remark.* When $X$ is not $\sigma$-finite and $p = 1$, the result in Theorem 5.6.11 does not hold necessarily: we refer to Exercises 5.6.9 and 5.6.10. On the other hand, if we consider $\mathbb{R}$ with the counting measure it is not $\sigma$-finite, but $\ell^1(\mathbb{R})^* = \ell^\infty(\mathbb{R})$ with the same proof than Proposition 5.6.3 (Exercise 5.6.11). The exact requirement for $L^1(X,\mu)^* \simeq L^\infty(X,\mu)$ is that $(X,\mu)$ is **localizable**; this means that every measurable set that is not a nullset contains a measurable subset of positive finite measure, and that every family of measurable sets admits (up to a nullset) a smallest measurable envelope. See Exercises 5.6.12 and 5.6.13 for details.

*5.6.13. Remark.* When $(X,\mathcal{A},\mu)$ is $\sigma$-finite and complete, the analog of Proposition 5.6.7 holds. That is, the dual of $L^\infty(X,\mathcal{A},\mu)$ can be identified with the space of complex finitely additive measures on $\mathcal{A}$ that are absolutely continuous with respect to $\mu$, with the total variation norm.

*5.6.14. Remark.* We will write the assignment from equation (5.17) as $f \mapsto \langle f, g\rangle$. Note that in the case $p = q = 2$ (i.e. Hilbert space), we have the actual inner product. In this notation, Hölder's inequality is written as $|\langle f, g\rangle| \leq \|f\|_p \, \|g\|_q$.

Theorem 5.6.11 gives us another proof that $L^p$ spaces are always Banach, as we have shown that for $p \in (1,\infty]$ we always have that $L^p(X)$ is the dual of $L^q(X)$, and duals of normed spaces are always Banach (Proposition 5.5.8). But we cannot use the same argument for $L^1(X)$ as it is not always a dual (it is in some cases, for instance $\ell^1(\mathbb{N}) = c_0^*$; but $L^1[0,1]$ with Lebesgue measure is not a dual, see Proposition 7.5.13). The general result that $L^p(X)$ is complete for $1 \leq p \leq \infty$ holds nonetheless, as we saw in Theorem 2.8.12.

**5.6.6. $C_0(T)$.** If we look at Theorem 2.9.2, when $\varphi \in C_0(T)^*$ and $\varphi(f) \geq 0$ whenever $f \geq 0$, then $\varphi$ is given by integration against a measure (note that $C_c(T)$ is dense in $C_0(T)$, so the dual is the same). It feels like this is really close to characterizing the dual! But we need to deal with arbitrary functionals, not just the positive ones, and that will require a bit of work.

As we will be discussing the complex version of Theorem 2.9.2, it could be argued that such result belongs in Section 2.10. But we had not discussed linear functionals nor boundedness then, so we feel it is more appropriate to include the theorem here. The techniques in the proof, though, all belong in Chapter 2.

**5.6.15. Theorem. (Riesz–Markov)**

*Let $T$ be locally compact Hausdorff, and $\varphi \in C_0(T)^*$. Then there exists a unique regular complex Borel measure $\gamma$ such that*

$$\varphi(f) = \int_T f\,d\gamma, \qquad f \in C_0(T). \tag{5.20}$$

*Also, $\|\varphi\| = |\gamma|(T)$. So $C_0(T)^*$ is isometrically isomorphic to the Banach space of regular complex Borel measures on $T$, via the duality (5.20).*

*Proof.* We assume that $\|\varphi\| = 1$, as then the general case will only require dealing with a scalar multiple. The idea of the proof is to use Theorem 2.9.2 on a well-chosen positive functional. Given $f \in C_c(T)$ with $f \geq 0$, we define

$$\psi(f) = \sup\{|\varphi(h)| : \ h \in C_c(T),\ |h| \leq f\}.$$

By definition $\psi(f) \geq 0$, and $\psi(cf) = c\psi(f)$ for all $c \in \mathbb{R}_+$. Next we want to show that $\psi$ is additive on $C_c(T)^+$. We have to show that for $f, g \geq 0$ we have $\psi(f+g) \leq \psi(f) + \psi(g)$ and $\psi(f+g) \geq \psi(f) + \psi(g)$. If $h \in C_c(T)$ with $|h| \leq f + g$, consider the open set $T_0 = \{f + g > 0\}$, and define

$$h_1 = \begin{cases} \dfrac{fh}{f+g}, & \text{in } T_0 \\ 0, & \text{in } T \setminus T_0 \end{cases}, \qquad h_2 = \begin{cases} \dfrac{gh}{f+g}, & \text{in } T_0 \\ 0, & \text{in } T \setminus T_0 \end{cases}$$

Noting that $|h_j| \leq |h|$, $j = 1, 2$, that $g + f = 0$ on $\partial T_0$ (since $T_0$ is open), and that $h_1, h_2$ are continuous on $T_0$ and have compact support, we conclude that $h_1, h_2 \in C_c(T)$. We also have $h_1 + h_2 = h$, $|h_1| \leq f$, $|h_2| \leq g$ (since $|h| \leq f + g$). Then

$$|\varphi(h)| = |\varphi(h_1 + h_2)| \leq |\varphi(h_1)| + |\varphi(h_2)| \leq \psi(f) + \psi(g).$$

As this can be done for any such $h$, we have shown that $\psi(f+g) \leq \psi(f)+\psi(g)$.

For the reverse inequality, fix $\varepsilon > 0$ and choose $h_1, h_2 \in C_c(T)$ with $\psi(f) \leq |\varphi(h_1)| + \varepsilon$ and $\psi(g) \leq |\varphi(h_2)| + \varepsilon$. Then, with $\varphi(h_j) = e^{i\theta_j}|\varphi(h_j)|$, $j = 1, 2$,

$$\begin{aligned} \psi(f) + \psi(g) &\leq |\varphi(h_1)| + |\varphi(h_2)| + 2\varepsilon = \varphi(e^{-i\theta_1}h_1 + e^{-i\theta_2}h_2) + 2\varepsilon \\ &\leq \psi(|e^{-i\theta_1}h_1 + e^{-i\theta_2}h_2|) + 2\varepsilon \leq \psi(|h_1| + |h_2|) + 2\varepsilon \\ &\leq \psi(f+g) + 2\varepsilon. \end{aligned}$$

As $\varepsilon > 0$ was arbitrary, we get the reverse inequality and $\psi(f+g) = \psi(f) + \psi(g)$.

For $f \in C_c(T)$ with $f = f_1 + if_2$ and $f_1, f_2$ real-valued, we define

$$\psi(f) = \psi(f_1^+) - \psi(f_1^-) + i\psi(f_2^+) - i\psi(f_2^-).$$

This definition does not depend on the actual $f_1^+, f_1^-$, in the sense that if $f = g_1 - g_2$ with $g_1, g_2 \geq 0$, then $\psi(g_1) - \psi(g_2) = \psi(f_1^+) - \psi(f_1^-)$, since $g_1 + f_1^- = f_1^+ + g_1$ is an equality of sums positive functions and the additivity of $\psi$ applies. From there it is routine to deduce that $\psi$ is linear in $C_c(T)$, as we did in the proof of Corollary 2.5.6.

With $\psi$ constructed, we note that it is positive by definition, and that $|\varphi(f)| \leq \psi(|f|) \leq \|f\|_\infty$. Indeed, if $f$ is real

$$\begin{aligned}|\varphi(f)| &= |\varphi(f^+) - \varphi(f^-)| \leq |\varphi(f^+)| + |\varphi(f^-)| \\ &\leq \psi(f^+) + \psi(f^-) = \psi(f^+ + f^-) = \psi(|f|).\end{aligned}$$

For arbitrary $f \in C_c(T)$,

$$\begin{aligned}|\varphi(f)| &= \big|\varphi(\operatorname{Re} f) + i\varphi(\operatorname{Im} f)\big| \leq |\varphi(\operatorname{Re} f)| + |\varphi(\operatorname{Im} f)| \\ &\leq \psi(|\operatorname{Re} f|) + \psi(|\operatorname{Im} f|) = \psi(|\operatorname{Re} f| + |\operatorname{Im} f|) \leq \sqrt{2}\,\psi(|f|).\end{aligned}$$

The coefficient $\sqrt{2}$ is not really necessary, as we will see below, but it is enough for now. Also, $\psi(|f|) \leq \|f\|_\infty$, since by definition $|\varphi(h)| \leq \|\varphi\|\,\|h\|_\infty \leq \|f\|_\infty$ for all $|h| \leq f$.

By Theorem 2.9.2 applied to the positive functional $\psi$, there exists a unique measure $\mu$ such that

$$\psi(f) = \int_T f\,d\mu, \qquad f \in C_c(T).$$

This $\mu$ is finite on compact sets, outer regular, inner regular on open and finite-measure sets, and complete. We have $\mu(T) = \psi(1) \leq \|1\|_\infty = 1$, so $\mu$ is finite and hence inner regular (Exercise 2.3.20). Now we have

$$|\varphi(f)| \leq \sqrt{2}\,\psi(|f|) = \sqrt{2}\int_T |f|\,d\mu = \sqrt{2}\,\|f\|_1, \qquad f \in C_c(T).$$

So $\varphi$ is bounded on a dense subspace of $L^1(T,\mu)$ (Proposition 2.8.18), hence it extends to $L^1(T,\mu)$ by Exercise 5.5.11. That is, $\varphi \in L^1(T,\mu)^*$. By Theorem 5.6.11 (we are dealing with a finite measure here), there exists $g \in L^\infty(T,\mu)$ with

$$\varphi(f) = \int_T f\,g\,d\mu, \qquad f \in L^1(T,\mu).$$

The equality holds in particular for all $f \in C_0(T)$. If we define $\gamma(E) = \displaystyle\int_E g\,d\mu$, then $\gamma$ is a regular complex Borel measure and

$$\varphi(f) = \int_T f\,d\gamma, \qquad f \in C_0(T)$$

(Exercises 2.10.9 and 2.10.17). As $\|\varphi\| = 1$, we have

$$\left|\int_T fg\,d\mu\right| = |\varphi(f)| \leq \|f\|_1$$

for all $f \in L^(T, \mu)$. Then $\|g\|_\infty \le 1$ by Proposition 5.6.8. Now

$$|\varphi(f)| = \left| \int_T f\, g\, d\mu \right| \le \int_T |f|\, d\mu = \psi(|f|),$$

showing that we can avoid the coefficient $\sqrt{2}$ that we had above. We also have

$$1 = \sup \big\{ |\varphi(f)| : \ f \in C_0(T),\ \|f\|_\infty = 1 \big\} \le \int_T |g|\, d\mu.$$

Since $|g| \le 1$ a.e. and $\mu(T) \le 1$, this forces $|g| = 1$ a.e. and $\mu(T) = 1$. The fact that $|g| = 1$ a.e. and Proposition 2.10.12, give us $|\gamma| = \mu$. Then

$$\|\varphi\| = 1 = \mu(T) = |\gamma|(T) = \|\gamma\|.$$

As for the uniqueness, if we had two measures realizing the same functional their difference would be—after checking that the difference of regular measures is regular—a regular complex Borel measure $\gamma$ with $\displaystyle\int_T f\, d\gamma = 0$ for all $f \in C_0(T)$. By Proposition 2.10.11 there exists a Borel function $h$ with $|h| = 1$ and $d\gamma = h\, d|\gamma|$. For any $f \in L^\infty(T, |\gamma|)$,

$$\left| \int_T hf\, d|\gamma| \right| = \left| \int_T f\, d\gamma \right| = |\varphi(f)| \le \|f\|_\infty.$$

By Proposition 5.6.8, $h \in L^1(T, |\gamma|)$. By Proposition 2.8.18 there exists a sequence $\{h_n\} \subset C_0(T)$ with $\|\bar{h} - h_n\|_1 \to 0$. Then

$$|\gamma|(X) = \int_T 1\, d|\gamma| = \int_T \bar{h}\, h\, d|\gamma| = \lim_n \int_T h_n\, h\, d|\gamma| = \lim_n \int_T h_n\, d\gamma = 0.$$

So $\gamma = 0$, showing the uniqueness. ☐

**5.6.7. $L^p[0,1]$, $0 \le p < 1$.** In this situation, the “$p$-norm” is not a norm (Exercise 5.6.14). One can still define

$$d_p(f, g) = \int_{[0,1]} |f - g|^p\, dm.$$

The condition $0 < p < 1$ guarantees that $d_p$ is a metric (from $|s+t|^p \le |s|^p + |t|^p$). And by definition this metric is translation invariant. Using the metric $d_p$ instead of the $p$-norm, the proof of Theorem 2.8.12 applies, and we get that $L^p$ is still a complete metric space. But things go downhill fast.

***5.6.16. Lemma.***

*Let $0 < p < 1$, and $\mathcal{X} = L^p[0,1]$, with Lebesgue measure. If $V \subset \mathcal{X}$ is nonempty, open, and convex, then $V = \mathcal{X}$.*

*Proof.* By translating, we may assume that $0 \in V$. As $V$ is open, there exists $\delta > 0$ such that $B_\delta(0) \subset V$.

Fix $f \in L^p[0,1]$. Choose $n \in \mathbb{N}$ such that $n^{p-1}\, d_p(f,0) < \delta$ (the fact that $p < 1$ guarantees that such choice is possible). Since the function $s \longmapsto \int_0^s |f|^p$ is continuous (Exercise 2.5.8 or Exercise 2.10.5) with range $[0, d_p(f,0)]$, for each $k = 1, \ldots, n$ there exists $s_k$ with $\int_0^{s_k} |f|^p = \frac{k}{n}\, d_p(f,0)$. For each $k$, define

$$g_k(t) = n\, f(t)\, 1_{[s_{k-1}, s_k]}.$$

We have $g_k \in L^p[0,1]$ and, due to the choice of the $s_k$,

$$d_p(g_k, 0) = \int_{s_{k-1}}^{s_k} n^p\, |f|^p = n^p\, \frac{d_p(f,0)}{n} = n^{p-1}\, d_p(f,0) < \delta.$$

So $g_k \in B_\delta(0) \subset V$ for all $k$, and by convexity

$$f = \sum_{k=1}^{n} \frac{1}{n}\, g_k \in V.$$

Thus $V = \mathcal{X}$. □

A striking consequence of the above is the following.

**5.6.17. *Proposition.***

*Let $0 < p < 1$, and $\mathcal{X} = L^p[0,1]$, with Lebesgue measure. Then $\mathcal{X}^* = \{0\}$.*

*Proof.* Let $\phi \in \mathcal{X}^*$. As $\phi$ is continuous and linear, $\phi^{-1}(B_1(0))$ is nonempty, open, and convex. By Lemma 5.6.16, $\phi^{-1}(B_1(0)) = \mathcal{X}$. So $|\phi(x)| < 1$ for all $x \in \mathcal{X}$, which implies $\phi = 0$. □

For future reference, we establish the following:

**5.6.18. *Proposition.***

*Let $X$ be a locally compact space and $\mu$ a Radon measure on $X$. Let $p \in (0,1)$. Then $C_c(X)$ is dense in $L^p(X, B(X), \mu)$.*

*Proof.* The proofs of Propositions 2.8.14, 2.8.16 and 2.8.18 actually work the same for $0 < p < 1$. □

**5.6.8. $\ell^p(\mathbb{N})$, $0 < p < 1$.** These spaces provide examples of topological vector spaces—metric, in fact—where the dual is big enough to separate points, but not big enough to have Hahn–Banach separation.

Again here we consider the topology given by the metric $d_p(x,y) = \sum_n |x_n - y_n|^p$. As opposed to the continuous case, the dual is "big enough" in some sense:

**5.6.19. Proposition.**

*Let $p \in (0,1)$. Then $\ell^p(\mathbb{N})^*$ separates points in $\ell^p(\mathbb{N})$.*

*Proof.* For each $n$, let $\varphi_n(x) = x_n$. Since convergence in the $d_p$ metric implies pointwise convergence, each $\varphi_n$ is continuous, so $\varphi_n \in \ell^p(\mathbb{N})^*$. Given $x \neq y$, there exists $m \in \mathbb{N}$ with $x_m \neq y_m$. Then $\varphi_m(x) \neq \varphi_m(y)$, so the dual separates points. $\square$

We can easily find more elements from the dual. For any $y \in \ell^\infty(\mathbb{N})$ and $x \in \ell^p(\mathbb{N})$, we have

$$\sum_k |x_k y_k|^p \leq \|y\|_\infty^p \sum_k |x_k|^p \leq \|y\|_\infty^p \, d_p(x,0).$$

So the linear functional $\varphi_y : x \longmapsto \sum_k x_k y_k$ is continuous. Conversely, since $\ell^p(\mathbb{N}) \subset \ell^1(\mathbb{N})$ is dense in the $\ell^1(\mathbb{N})$ norm (already $c_{00}$ is dense in $\ell^1(\mathbb{N})$ and it lies inside $\ell^p(\mathbb{N})$), the dual of $\ell^p(\mathbb{N})$ can be identified with $\ell^\infty(\mathbb{N})$, though not in a topological way. We say "not topological", because the identification occurs using the $\|\cdot\|_1$ and not $\|\cdot\|_p$.

Things go wrong here in more than one way, as shown by the next two results. Namely, $\ell^p(\mathbb{N})$ is not locally convex (Proposition 5.6.20) and, more importantly, in language to be established in Section 5.7, there is no Hahn–Banach theorem (Proposition 5.6.21).

**5.6.20. Proposition.**

*Let $p \in (0,1)$. Then $\ell^p(\mathbb{N})$ is not locally convex.*

*Proof.* Suppose that it is locally convex. Then the unit ball $B_1(0)$ contains an open convex neighbourhood $V$ with $0 \in V$. Since $V$ is an open neighbourhood of 0, there exists $r > 0$ with $B_r(0) \subset V \subset B_1(0)$. Fix $s > 0$ with $s < r$. Since $s^{1/p} e_j \in B_r(0)$ for all $j$, by the convexity of $V$ we get that, for all $m \in \mathbb{N}$,

$$\sum_{j=1}^m \frac{s^{1/p}}{m} e_j \in V.$$

Then

$$1 > d\Big(\sum_{j=1}^m \frac{s^{1/p}}{m} e_j, 0\Big) = \sum_{j=1}^m \frac{s}{m^p} = s\, m^{1-p}.$$

As $1-p>0$ and $s>0$ and we can do this for all $m\in\mathbb{N}$, we obtain a contradiction. $\square$

**5.6.21. Proposition.** ([Pec65])

*Let $p\in(0,1)$. There exists a closed subspace $M\subset\ell^p(\mathbb{N})$ such that $d(e_1,M)=\frac{1}{2}$ while $\varphi(e_1)=0$ for all $\varphi\in\ell^p(\mathbb{N})^*$ such that $\varphi|_M=0$. That is, it is not possible to separate $M$ and $e_1$.*

*Proof.* Fix $n\in\mathbb{N}$. Define a sequence $b=\{k^{-1/p}\}_k$. We have

$$\sum_k |b_k| = \sum_k \frac{1}{k^{1/p}} < \infty, \qquad \sum_k |b_k|^p = \sum_k \frac{1}{k} = \infty$$

So $b\in\ell^1(\mathbb{N})\setminus\ell^p(\mathbb{N})$. Let $s_n\in\mathbb{N}$ be such that

$$\sum_{k=1}^{s_n} b_k^p \geq n^p\|b\|_1^p.$$

Define numbers

$$a_{n,k} = \frac{b_k}{\left(\sum_{j=1}^{s_n} b_j^p\right)^{1/p}}$$

Then

$$\sum_{k=1}^{s_n} a_{n,k} \leq \frac{\sum_{k=1}^{s_n} b_k}{\left(\sum_{j=1}^{s_n} b_j^p\right)^{1/p}} \leq \frac{\|b\|_1}{n\|b\|_1} = \frac{1}{n},$$

and

$$\sum_{k=1}^{s_n} a_{n,k}^p = 1.$$

Let $\mathbb{N}\setminus\{1\}=\bigcup_n S_n$, with the union disjoint and $|S_n|=s_n$. Write $S_n=\{c_{n,1},\dots,c_{n,s_n}\}$ and define

$$x_n = \sum_{k=1}^{s_n} a_{n,k} e_{c_{n,k}}.$$

Let

$$M=\overline{\operatorname{span}}\{e_1+x_n:\ n\in\mathbb{N}\}.$$

We have

$$\|e_1-(e_1+x_n)\|_1 = \|x_n\|_1 = \sum_{k=1}^{s_n} a_{n,k} \leq \frac{1}{n},$$

so $e_1$ is in the $\|\cdot\|_1$-closure of $M$. Thus, if $\varphi \in \ell^p(\mathbb{N})^*$ with $\varphi|_M = 0$, as $\varphi$ is implemented by some $y \in \ell^\infty(\mathbb{N})$ we get that $\varphi$ is $\|\cdot\|_1$-continuous and then

$$|\varphi(e_1)| = \lim_n \varphi(e_1 + x_n) = 0.$$

So $e_1$ cannot be separated from $M$ by any element in $\ell^p(\mathbb{N})^*$. But, for any linear combination $\sum_{n=1}^m \beta_n(e_1 + x_n)$, since the sets $S_n$ are disjoint,

$$\begin{aligned} d\big(e_1, \sum_{n=1}^m \beta_n(e_1+x_n)\big) &= \sum_{k=1}^\infty \Big| e_1(k) - \sum_{n=1}^m \beta_n(e_1(k) + x_n(k))\Big|^p \\ &= \big|1 - \sum_{n=1}^m \beta_n\big|^p + \sum_{k=2}^\infty \Big|\sum_{n=1}^m \beta_n\, x_n(k)\Big|^p \\ &= \big|1 - \sum_{n=1}^m \beta_n\big|^p + \sum_{h=1}^\infty \sum_{k\in S_h} \Big|\sum_{n=1}^m \beta_n\, x_n(k)\Big|^p \\ &= \big|1 - \sum_{n=1}^m \beta_n\big|^p + \sum_{h=1}^m \sum_{k\in S_h} |\beta_h|^p\, |a_{h,c_h^{-1}(k)}|^p \\ &= \big|1 - \sum_{n=1}^m \beta_n\big|^p + \sum_{h=1}^m |\beta_h|^p \geq \frac{1}{2}. \end{aligned}$$

The last inequality is clear if $\sum_{n=1}^m |\beta_n|^p \geq \frac{1}{2}$; and if $\sum_{n=1}^m |\beta_n|^p < \frac{1}{2}$, then $\sum_{n=1}^m |\beta_n| \leq \frac{1}{2}$ and

$$\big|1 - \sum_{n=1}^m \beta_n\big| \geq 1 - \big|\sum_{n=1}^m \beta_n\big| \geq 1 - \sum_{n=1}^m |\beta_n| \geq \frac{1}{2}.$$

Thus $d(e_1, M) \geq \frac{1}{2}$. $\square$

*(The following comment is out of place, in the sense that weak topologies have not been defined yet. But a more advanced reader who came back to look for the examples, might find it useful.)* It is interesting to notice that, if we consider $\ell^p(\mathbb{N})$ with its weak topology, by Proposition 7.1.5 the dual is still the same. As the weak topology is locally convex, now Hahn–Banach applies. This is no contradiction, as now the closed sets are not the same. In particular, the subspace $M$ in Proposition 5.6.21 is not weakly closed—as explicitly shown in the proof.

**5.6.9. Exercises.**

(5.6.1) Let $q \in [1,\infty)$, $p \in \mathbb{R}$ with $\frac{1}{p}+\frac{1}{q}=1$, $g:\mathbb{N}\to\mathbb{C}$. Put $B_p=\{f \in \ell^p(\mathbb{N}) :\ \|f\|_p=1\}$. Show that

$$\max\{\,|\langle f,g\rangle| :\ f\in B_p\}=\max\left\{\sum_{k=1}^{\infty}|g(k)f(k)| :\ f\in B_p\right\},$$

where

$$\langle f,g\rangle=\sum_{k=1}^{\infty}g(k)f(k).$$

(5.6.2) Consider the Banach space $c_0$ (Example 5.1.9). Given $x\in c_0$ show that for any $f\in\ell^1(\mathbb{N})$ the map $x\mapsto\sum_n x_nf_n$ is a continuous linear functional. Use this to prove that there is an isometric embedding of $\ell^1(\mathbb{N})$ into $c_0^*$. Prove that this embedding is surjective, i.e. $c_0^*=\ell^1(\mathbb{N})$.

(5.6.3) Using the ideas in Exercise 5.6.2, show that the dual of $c$ (the Banach space of convergent sequences with the supremum norm) is $\ell^1(\mathbb{N})$.

(5.6.4) Show that $c$ and $c_0$ are isomorphic as Banach spaces.

(5.6.5) Show that there is no isometry—linear or not—between $c$ and $c_0$ *(Hint: consider that 1 is the middle point between 0 and 2)*.

(5.6.6) Let $S$ be an arbitrary set. Show that $\ell^1(S)^*=\ell^\infty(S)$, where the duality is the same as the one in Exercise 5.6.2.

(5.6.7) Let $p\in(1,\infty)$ and $g\in\ell^q(\mathbb{N})$. Show that the map

$$\varphi : f\mapsto\sum_{k=1}^{\infty}f(k)g(k) \tag{5.21}$$

defines a bounded functional on $\ell^p(\mathbb{N})$ with norm $\|g\|_q$

(5.6.8) Let $\varphi\in\ell^\infty(\mathbb{N})^*$. Show that there exist $\varphi_1,\varphi_\infty$ such that $\varphi=\varphi_1+\varphi_\infty$, where $\varphi_1,\varphi_\infty\in\ell^\infty(\mathbb{N})^*$, $\varphi_\infty|_{c_0}=0$, and $\varphi_1(x)=\langle x,y\rangle$ for some $y\in\ell^1(\mathbb{N})$ and all $x\in\ell^\infty(\mathbb{N})$.

(5.6.9) Let $\mathcal{M}$ be the $\sigma$-algebra of subsets of $[0,1]$ that are either finite or countable, or alternatively have finite or countable complement. Let $\mu$ be the counting measure.

(i) Show that $\mu$ is not $\sigma$-finite.

(ii) Show that $g(x)=x$ is not measurable.

(iii) Show that $\gamma : f\longmapsto\sum_x xf(x)$ defines a bounded linear functional on $L^1(\mu)$.

(iv) Show that $\gamma$ is not of the form $\gamma(f)=\int fh\,d\mu$ for $h\in L^\infty(\mu)$.

(v) Conclude that $L^1(\mu)^*\neq L^\infty(\mu)$.

(5.6.10) Let $X=\{0,1\}$, and $\mu$ the measure given by $\mu(\{0\})=1$, $\mu(\{1\})=\infty$. Show that $L^1(X)^*\neq L^\infty(X)$.

(5.6.11) We consider the measure space $[0,1]$ with the counting measure.

(i) Show that $\ell^1[0,1]^* = \ell^\infty[0,1]$, though the measure is not $\sigma$-finite.

(ii) Let $\mathcal{X} = \{a : [0,1] \to \mathbb{C} : \operatorname{supp} a \text{ is countable and } a \in C_0(\operatorname{supp} a)\}$, where the topology is given by the $\|\cdot\|_\infty$ norm . Show that $\mathcal{X}$ is complete and that $\mathcal{X}^* = \ell^1[0,1]$.

(5.6.12) Recall from Exercise 2.3.26 that a measure $\mu$ is **semifinite** if for every measurable $E$ with $\mu(E) = \infty$ there exists $F \subset E$ with $0 < \mu(F) < \infty$. Show that for a measure space $(X, \mathcal{A}, \mu)$ the following statements are equivalent:

(i) $\mu$ is seminifite;

(ii) the canonical embedding $\Gamma : L^\infty(X) \to L^1(X)^*$ is injective.

When these conditions are satisfied, $\Gamma$ is actually isometric.

(5.6.13) *(This exercise is non-trivial, and it appears as an exercise by necessity of space; if tackled, it should be considered a project, and some guidance will likely be needed)* A measure $\mu$ on $(X, \mathcal{A})$ is **localizable** if it is semifinite and, in addition, given any collection $\mathcal{E}$ of measurable sets, it admits an **essential supremum**: that is a measurable $H$ such that $\mu(E \setminus H) = 0$ for all $E \in \mathcal{E}$ (so $E \subset H$ a.e.) and if $H'$ satisfies the same property then $\mu(H \setminus H') = 0$ (that is, $H$ is the smallest such set up to nullsets).

(i) Show that if $\mu$ is semifinite and the canonical map $\Gamma : L^\infty(X) \to L^1(X)^*$ is surjective, then $\mu$ is localizable.

(ii) *(This part of the exercise is pure measure theory, but it is needed for the rest; it allows us to patch measurable functions—notably Radon-Nikodym derivatives—together as long as they agree almost everywhere on the intersection of their domains. It is not a trivial result so the reader might want to skip it and just use it)* Suppose that $\mu$ is localizable and that $\{f_j\}$ is a family of measurable real-valued functions, each with domain $D_j \in \mathcal{A}$ and such that $f_j = f_k$ a.e. on $D_j \cap D_k$. For each $q \in \mathbb{Q}$ and each $j$, let
$$E_{j,q} = \{x \in D_j : \ f_j(x) \geq q\},$$
and let $E_q$ be an essential supremum of $\{E_{j,q} : \ j\}$. Put
$$h'(x) = \sup\{q : \ q \in \mathbb{Q}, \ x \in E_q\},$$
allowing for $\sup \varnothing = -\infty$. Finally, let $h(x) = h'(x)$ if $h'(x) \in \mathbb{R}$ and zero otherwise. Show that $h$ is measurable and that $h|_{D_j} = f_j$ a.e. for all $j$.

(iii) Show that $\Gamma$ is an isometric isomorphism if and only if $\mu$ is localizable.

(5.6.14) For $0 < p < 1$, show that $\|f\|_p = \left(\sum_j |f(j)|^p\right)^{1/p}$ is not a norm on $\ell^p(\mathbb{N})$.

(5.6.15) Show that if $0 < p < 1$ and $q$ satisfies $\frac{1}{p} + \frac{1}{q} = 1$, we have the **Reverse Hölder Inequality**: given $f \in \ell^p(\mathbb{N})$, $g \in \ell^q(\mathbb{N})$,

$$\sum_k |f(k)g(k)| \geq \|f\|_p \|g\|_q.$$

(5.6.16) Prove that $\psi : L^q(X) \to L^p(X)^*$, as defined in (5.17), is linear and bounded, and $\|\psi(g)\| \leq \|g\|_q$.

(5.6.17) Fix $p \in [1, \infty)$. Let $\mathcal{X} = C[0,1]$ seen as a normed space with the $p$-norm. Show that the functional $f \longmapsto f(0)$ is unbounded.

(5.6.18) For a fixed $0 < p < 1$, consider the vector space

$$\ell^p(\mathbb{N}) = \{x : \mathbb{N} \to \mathbb{C} : \sum_j |x_j|^p < \infty\},$$

with $d_p(x,y) = \sum_j |x_j - y_j|^p$.

(i) Show that $d_p$ is a metric, so $\ell^p(\mathbb{N})$ is a TVS.

(ii) Show that the dual of $\ell^p(\mathbb{N})$ is $\ell^\infty(\mathbb{N})$.

(5.6.19) *(another example of a topological vector space with trivial dual)* Let $\mathcal{X} = B[0,1]$, the bounded Borel functions modulo almost everywhere equality. On this complex vector space, define

$$d(f,g) = \int_0^1 \frac{|f-g|}{1+|f-g|}.$$

(i) Show that $d$ is a distance.

(ii) Show that $(\mathcal{X}, d)$ is a topological vector space.

(iii) Show that $\mathcal{X}^* = \{0\}$.

## 5.7. The Hahn–Banach Theorem

The Hahn–Banach Theorem is arguably the most important result in Functional Analysis. It opens many doors in working with infinite-dimensional spaces and their operators. The name applies to two different kinds of incarnations of the result, the first one about extending bounded linear functionals, and the second one about separating disjoint convex subsets. We will obtain great mileage from both.

In the proof of Theorem 5.7.5 we will use the fact that a complex vector space can also be seen as a real vector space. Indeed, if $V$ is a complex vector space with basis $\{f_j\}$, then $\{f_j\} \cup \{i\, f_j\}$ is a basis of $V$ as a real vector space.

In parts of this section we will use $V, W, Z$ instead of $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ to denote our vector spaces, to emphasize that in those cases we are not really using any underlying topology and we are just working on the algebraic vector spaces.

We remark the fact that by "real functional" we mean a functional that takes real values and, moreover, is linear only for real scalars—note that "real valued and complex linear" makes no sense.

**5.7.1. Lemma.**

*Let $V$ be a complex vector space and $\psi : V \to \mathbb{R}$. Then the following statements are equivalent:*

*(i) there exists $\varphi : V \to \mathbb{C}$ linear, with $\psi(x) = \operatorname{Re}\varphi(x)$ for all $x \in V$;*

*(ii) $\psi$ is real linear.*

*When the statements hold, we have the relation $\varphi(x) = \psi(x) - i\,\psi(ix)$ for all $x \in V$*

*Proof.* For any complex number $\omega$, we have the equality $\omega = \operatorname{Re}\omega - i\operatorname{Re}(i\omega)$. This is the key to the proof.

(i) $\implies$ (ii) immediate from the linearity of $\varphi$ and the real linearity of $\operatorname{Re}$.

(ii) $\implies$ (i). If $\psi$ is a real linear functional and we let $\varphi(x) = \psi(x) - i\,\psi(ix)$, then clearly $\psi = \operatorname{Re}\varphi$ and $\varphi(x+y) = \varphi(x) + \varphi(y)$. What is less obvious is the behaviour with complex scalars. But for $a, b \in \mathbb{R}$

$$\begin{aligned}
\varphi((a+ib)x) &= \varphi(ax + ibx) = \varphi(ax) + \varphi(bix) = a\varphi(x) + b\varphi(ix) \\
&= a\psi(x) - a\,i\,\psi(ix) + b\psi(ix) - b\,i\,\psi(-x) \\
&= a\psi(x) + ib\psi(x) - a\,i\,\psi(ix) - ib\,i\,\psi(ix) \\
&= (a+ib)(\psi(x) - i\psi(ix)) = (a+ib)\varphi(x).
\end{aligned}$$

□

**5.7.2. Definition.**

*A **real seminorm** is a function $q : \mathcal{X} \to [0, \infty)$ such that*

*(i) $q(tx) = |t|\, q(x)$ for all $t \in \mathbb{R}$ (**homogeneity**);*

*(ii) $q(x+y) \le q(x) + q(y)$ (**subadditivity/triangle inequality**).*

A sublinear function (Definition 5.4.8) is like a real seminorm but without the requirement $q(-x) = q(x)$, and in principle $q$ can take negative values.

The next lemma contains the key step to prove Hahn–Banach's Theorem (5.7.5).

**5.7.3. Lemma.**

*Let $V$ be a real vector space, $Z \subset V$ a subspace, and $v_0 \in V \setminus Z$. Suppose that $\psi : Z \to \mathbb{R}$ is a real linear functional, and there exists a sublinear function $q : V \to \mathbb{R}$ such that $\psi(x) \leq q(x)$ for all $x \in Z$. Then there exists an extension $\tilde{\psi} : Z + \mathbb{R}v_0 \to \mathbb{R}$ of $\psi$ with $\tilde{\psi}(x) \leq q(x)$ for all $x \in Z + \mathbb{R}v_0$.*

*Proof.* Since $v_0$ and $Z$ are linearly independent, $Z + \mathbb{R}v_0 = \{cv_0 + z : \ c \in \mathbb{R},\ z \in Z\}$, with a unique $c$ and $z$ for each element. Fix $z_1, z_2 \in Z$; then

$$\psi(z_1 + z_2) \leq q(z_1 + z_2) \leq q(z_1 + v_0) + q(z_2 - v_0),$$

and we deduce that

$$\psi(z_2) - q(z_2 - v_0) \leq -\psi(z_1) + q(z_1 + v_0)$$

This can be done for all $z_1, z_2 \in Z$, so we conclude that there exists $d \in \mathbb{R}$ with

$$\sup\{\psi(z_2) - q(v_0 - z_2) : z_2 \in Z\} \leq d \leq \inf\{-\psi(z_1) + q(v_0 + z_1) : \ z_1 \in Z\}.$$

Now define the linear map $\tilde{\psi}(cv_0 + z) = cd + \psi(z)$. Then $\tilde{\psi}|_Z = \psi$. Fix $c \in \mathbb{R} \setminus \{0\}$, $z \in Z$. Suppose first that $c > 0$. Then

$$\begin{aligned}
\tilde{\psi}(cv_0 + z) = cd + \psi(z) &= c\left(d + \psi\left(\frac{z}{c}\right)\right) \\
&\leq c\left(-\psi\left(\frac{z}{c}\right) + q\left(v_0 + \frac{z}{c}\right) + \psi\left(\frac{z}{c}\right)\right) \\
&= q(cv_0 + z).
\end{aligned}$$

Similarly, if $c < 0$,

$$\begin{aligned}
\tilde{\psi}(cv_0 + z) = cd + \psi(z) &= (-c)\left(-d + \psi\left(\frac{z}{-c}\right)\right) \\
&\leq (-c)\left(-\psi\left(\frac{z}{-c}\right) + q\left(\frac{z}{-c} - v_0\right) + \psi\left(\frac{z}{-c}\right)\right) \\
&= q(cv_0 + z).
\end{aligned}$$

□

The Hanh–Banach theorem, in its general form, requires no topology whatsoever. So while most of the time in functional analysis we care about closed subspaces, the subspaces in the general form of Hahn–Banach are just plain subspaces.

We prove first the sublinear version, which together with Lemma 5.7.3 contain all the core technicalities of the result. Then we will produce other versions based on this, that will be used many times and form a crucial family of results in functional analysis.

**5.7.4. Theorem. (The Hahn–Banach Theorem, sublinear version)**

*Let $V$ be a real vector space with a sublinear function $q$. Let $W$ be a subspace of $V$, and $\varphi : W \to \mathbb{R}$ a real linear functional that satisfies $\varphi(x) \leq q(x)$ for all $x \in W$. Then there exists a real linear functional $\tilde{\varphi} : V \to \mathbb{R}$ such that $\tilde{\varphi}|_W = \varphi$ and $\tilde{\varphi}(x) \leq q(x)$ for all $x \in V$.*

*Proof.* Let $\mathcal{F}$ be the family of all $(W', \varphi')$, where $W'$ is a subspace of $V$ with $W \subset W'$, $\varphi' : W' \to \mathbb{R}$ extends $\varphi$, and $\varphi'(x) \leq q(x)$ for all $x \in W'$; the family $\mathcal{F}$ is trivially nonempty as $(W, \varphi) \in \mathcal{F}$. In $\mathcal{F}$, we consider the partial order $(W_1, \varphi_1) \leq (W_2, \varphi_2)$ if $W_1 \subset W_2$ and $\varphi_2|_{W_1} = \varphi_1$. Let $\{(W_j, \varphi_j)\}$ be a chain in $\mathcal{F}$; put $W' = \cup_j W_j$, which being an increasing union of subspaces is a subspace of $V$, and let $\varphi' : W' \to \mathbb{R}$ be given by $\varphi'(x) = \varphi_j(x)$ if $x \in W_j$. The compatibility given by the order guarantees that $\varphi'$ is well-defined. It is clear that $(W', \varphi') \in \mathcal{F}$ and it is an upper bound for the chain. Then, by Zorn's Lemma there exists a maximal element $(Z, \tilde{\varphi})$ in $\mathcal{F}$. If $Z \subsetneq V$, we can use Lemma 5.7.3 to contradict the maximality of $(Z, \tilde{\varphi})$. So $Z = V$ and $\tilde{\varphi}$ is the desired extension. The condition $\tilde{\varphi} \leq q$ comes for free since every element in $\mathcal{F}$ satisfies it.

$\square$

**5.7.5. Theorem. (The Hahn–Banach Theorem)**

*Let $V$ be a complex vector space with a seminorm $q$. Let $W$ be a (not necessarily closed) subspace of $V$ and $\varphi : W \to \mathbb{C}$ a linear functional satisfying $|\varphi(x)| \leq q(x)$ for all $x \in W$. Then there exists a linear functional $\tilde{\varphi} : V \to \mathbb{C}$ with $\tilde{\varphi}|_W = \varphi$ and $|\tilde{\varphi}(x)| \leq q(x)$ for all $x \in V$.*

*The result holds with $\mathbb{R}$ in place of $\mathbb{C}$, $\varphi$ real linear, and $q$ a real seminorm.*

*Proof.* In the real case, we use Theorem 5.7.4 to get $\tilde{\varphi} : V \to \mathbb{R}$ with $\tilde{\varphi}|_W = \varphi$ and $\tilde{\varphi}(x) \leq q(x)$ for all $x \in V$. Because $q$ is a seminorm and thus $q(-x) = q(x)$, for any $x \in V$ we have the inequality $\tilde{\varphi}(-x) \leq q(-x) = q(x)$; this we may write as $-q(x) \leq \tilde{\varphi}(x)$. Together with $\tilde{\varphi}(x) \leq q(x)$, this gives us $|\tilde{\varphi}(x)| \leq q(x)$.

Now consider the complex case. Let $\varphi_1$ be the real linear functional given by $\varphi_1(x) = \operatorname{Re} \varphi(x)$. Then $|\varphi_1(x)| \leq |\varphi(x)| \leq q(x)$ for all $x \in W$, so we can apply the previous part of the proof to obtain $\tilde{\varphi}_1 : V \to \mathbb{R}$ with $\tilde{\varphi}_1|_W = \operatorname{Re} \varphi$ and $|\tilde{\varphi}_1(x)| \leq q(x)$ for all $x \in W$. Using Lemma 5.7.1, define a new functional $\tilde{\varphi}(x) = \tilde{\varphi}_1(x) - i\,\tilde{\varphi}_1(ix)$. Then, if $x \in W$, $\tilde{\varphi}(x) = \varphi_1(x) - i\varphi_1(ix) =$

$\operatorname{Re}\varphi(x) - i\operatorname{Re}\varphi(ix) = \varphi(x)$. Fix $x \in V$. Let $|\tilde{\varphi}(x)|\, e^{i\theta} = \tilde{\varphi}(x)$ be the polar form of the complex number $\tilde{\varphi}(x)$. Then $|\tilde{\varphi}(x)| = e^{-i\theta}\tilde{\varphi}(x) = \tilde{\varphi}(e^{-i\theta}x)$. Since $\operatorname{Re}\tilde{\varphi}(e^{-i\theta}x) \leq q(e^{i\theta}x)$,

$$|\tilde{\varphi}(x)| = \tilde{\varphi}(e^{-i\theta}x) = \operatorname{Re}\tilde{\varphi}(e^{-i\theta}x) \leq q(e^{-i\theta}x) = q(x). \qquad \square$$

The key feature of the Hahn–Banach Theorem is that we can extend functionals while still controlling the norm. The most common case we will use is that where the seminorm is actually a norm; but it is also interesting to note that the above formulation allows us to use Hahn–Banach in locally convex spaces (and we will!). The first corollary below is what would have been called "The Hahn–Banach Theorem" if we were just working with normed spaces.

***5.7.6. Corollary. (Hahn–Banach Theorem, normed spaces)***

*Let $\mathcal{X}$ be a normed space, $\mathcal{Y} \subset \mathcal{X}$ a (not necessarily closed) subspace, and $\varphi : \mathcal{Y} \to \mathbb{C}$ a bounded linear functional. Then there exists $\tilde{\varphi} : \mathcal{X} \to \mathbb{C}$ linear and bounded, with $\|\tilde{\varphi}\| = \|\varphi\|$.*

*Proof.* Apply Theorem 5.7.5 with $q(x) = \|\varphi\|\,\|x\|$. $\square$

Now we can prove that $\mathcal{X}^*$ is nontrivial for any normed space $\mathcal{X}$.

***5.7.7. Corollary.***

*Let $\mathcal{X}$ be a normed vector space. Then $\mathcal{X}^*$ separates points in $\mathcal{X}$ and, for any $x \in \mathcal{X}$,*

$$\|x\| = \max\{|\varphi(x)| : \ \varphi \in \mathcal{X}^* \text{ and } \|\varphi\| \leq 1\}. \tag{5.22}$$

*Proof.* Note that (5.22) implies that $\mathcal{X}^*$ separates points: if $\varphi(x) = \varphi(y)$ for all $\varphi \in \mathcal{X}^*$, then by (5.22) $\|x - y\| = 0$, i.e. $x = y$. So we need to prove (5.22).

Fix $x \in \mathcal{X}$; for any $\varphi \in \mathcal{X}^*$ with $\|\varphi\| \leq 1$,

$$|\varphi(x)| \leq \|\varphi\|\,\|x\| \leq \|x\|.$$

Thus the right-hand-side in (5.22) is less or equal than the left-hand-side. Now let $W = \mathbb{C}x$, the subspace spanned by $x$. Let $\varphi_0 : W \to \mathbb{C}$ be given by $\varphi_0(cx) = c\|x\|$. Note that $\varphi_0$ is linear and that $\|\varphi_0\| = 1$. By the Hahn–Banach Theorem (Corollary 5.7.6), there exists $\varphi \in \mathcal{X}^*$ with $\|\varphi\| = 1$ and $\varphi|_W = \varphi_0$. In particular, $\varphi(x) = \|x\|$. This establishes (5.22). $\square$

It is interesting to compare (5.22) with how we defined the norm in $\mathcal{X}^*$:

$$\|\varphi\| = \sup\{|\varphi(x)| : \ x \in \mathcal{X} \ \text{ and } \|x\| = 1\}.$$

This reinforces the idea that there is a "duality" between the roles of $x$ and $\varphi$ in the expression $\varphi(x)$. Such duality will appear more and more often as we move along.

The next result holds for locally convex spaces (Lemma 5.7.23), but we will need to wait until Corollary 5.7.19 to be able to prove it in its general form.

***5.7.8. Lemma.***

*Let $\mathcal{X}$ be an infinite-dimensional normed space. Then $\mathcal{X}^*$ is infinite-dimensional.*

*Proof.* Let $\{x_n\} \subset \mathcal{X}$ be linearly independent. For each $n$, let

$$\mathcal{X}_n = \operatorname{span}\{x_1, \ldots, x_n\}$$

and define $\varphi_n : \mathcal{X}_n \to \mathbb{C}$ with $\varphi_n(x_n) = 1$ and $\varphi_n(x_k) = 0$ for all $k < n$. By Corollary 5.2.5 $\varphi_n$ is bounded on $\mathcal{X}_n$, and so it admits a bounded extension—still denoted $\varphi_n$—to all of $\mathcal{X}$ by Corollary 5.7.6. Suppose that $\sum_{j=1}^m \alpha_j \varphi_j = 0$. Evaluating at $x_1$, we get $\alpha_1 = 0$. Now evaluating at $x_2$ we get $\alpha_2 = 0$. Continuing in the same way, we get $\alpha_1 = \ldots = \alpha_m = 0$, and so $\varphi_1, \ldots, \varphi_m$ are linearly independent. Note that we never needed to evaluate $\phi_j(x_k)$ with $k > j$. It follows that $\{\varphi_j\}$ are linearly independent, and so $\dim \mathcal{X}^* = \infty$. $\square$

Another way to show that $\varphi_n$ in Lemma 5.7.8 is bounded, before it is extended, is to use Theorem 5.2.2. Namely, we can define a norm by letting $\|x\|_{\varphi_n} = \|x\| + |\varphi_n(x)|$. Since $\dim \mathcal{X}_n < \infty$, by Theorem 5.2.2 there exists $c > 0$ such that $\|x\| + |\varphi_n(x)| \leq c\|x\|$ for all $x \in \mathcal{X}_n$. Then $|\varphi_n(x)| \leq (c-1)\|x\|$, and $\varphi_n$ is bounded.

*5.7.9. Remark*. In the case where $\mathcal{X}$ is a normed space, here is an alternative proof of Proposition 5.5.12. Suppose that $\bigcap_{j=1}^n \ker \varphi_j = \{0\}$. By Lemma 5.5.10, any $\varphi \in \mathcal{X}^*$ is a linear combination of $\varphi_1, \ldots, \varphi_n$. So $\mathcal{X}^*$ is finite-dimensional. By Lemma 5.7.8, it follows that $\mathcal{X}$ is finite-dimensional.

It is important to emphasize that Theorem 5.7.5 does not require the vector space to be a normed space; it doesn't even require a topology, just a fixed seminorm. As a consequence, the theorem can be somehow used in locally convex spaces. We will soon see very useful versions of Hahn–Banach that apply to locally convex spaces (Theorems 5.7.13 and 5.7.18).

The next lemma is a particular case of Corollary 5.7.19, but it is what we can prove with the available tools.

**5.7.10. Lemma.**

*Let $\mathcal{X}$ be a normed space, $\mathcal{Z} \subset \mathcal{X}$ a not-dense subspace, and $x_0 \in \mathcal{X} \setminus \overline{\mathcal{Z}}$. Then there exists a nonzero bounded (on $\overline{\mathcal{Z}} + \mathbb{C}x_0$) linear functional $\psi : \mathcal{Z} + \mathbb{C}x_0 \to \mathbb{C}$ such that $\psi|_{\overline{\mathcal{Z}}} = 0$.*

*Proof.* Define $\psi(z + cx_0) = c$. The only issue is whether this is bounded. Note that $\|cx_0 + z\| > 0$ for all $c \neq 0$ and all $z \in \mathcal{Z}$, because if it is zero then $x_0 = -c^{-1}z \in \mathcal{Z}$. Thus

$$\begin{aligned}
\|\psi\| &= \sup\left\{\frac{|c|}{\|cx_0 + z\|} : \ c \in \mathbb{C} \setminus \{0\}, \ z \in \mathcal{Z}\right\} \\
&= \sup\left\{\frac{1}{\|x_0 + z/c\|} : \ c \in \mathbb{C} \setminus \{0\}, \ z \in \mathcal{Z}\right\} \\
&= \sup\left\{\frac{1}{\|x_0 + z\|} : \quad z \in \mathcal{Z}\right\} = \sup\left\{\frac{1}{\|x_0 - z\|} : \quad z \in \mathcal{Z}\right\} \\
&= \frac{1}{\inf\{\|x_0 - z\| : \quad z \in Z\}} = \frac{1}{\operatorname{dist}(x_0, \mathcal{Z})}.
\end{aligned}$$

□

**5.7.11. Corollary.**

*Let $\mathcal{X}$ be normed. If $\mathcal{X}^*$ is separable, then $\mathcal{X}$ is separable.*

*Proof.* Let $\{\varphi_n\}$ be a countable dense subset of $\mathcal{X}^*$. For each $n$, choose $x_n \in \mathcal{X}$ with $\|x_n\| = 1$ and $|\varphi_n(x_n)| \geq \frac{1}{2}\|\varphi_n\|$.

Let $\mathcal{Y} = \operatorname{span}\{x_n : \ n \in \mathbb{N}\}$. If $\mathcal{Y}$ is not dense in $\mathcal{X}$, by Lemma 5.7.10 there exists nonzero bounded $\varphi_0$ on $\mathcal{Y} + \mathbb{C}x_0$ for some $x_0$, with $\varphi_0(x_0) \neq 0$ and $\varphi_0|_{\mathcal{Y}} = 0$. By Corollary 5.7.6 there exists an extension $\varphi : \mathcal{X} \to \mathbb{C}$ with $\|\varphi\| = \|\varphi_0\| > 0$, and $\varphi|_{\mathcal{Y}} = 0$ (this also follows from Corollary 5.7.19). Now

$$\begin{aligned}
\frac{1}{2}\|\varphi_n\| &\leq |\varphi_n(x_n)| = |\varphi_n(x_n) - \varphi(x_n)| \\
&= |(\varphi_n - \varphi)(x_n)| \leq \|\varphi_n - \varphi\|.
\end{aligned} \tag{5.23}$$

By the density of $\{\varphi_n\}$ in $\mathcal{X}^*$, there exists a subsequence $\{\varphi_{n_k}\}$ with $\|\varphi_{n_k} - \varphi\| \to 0$. But then (5.23) gives $\|\varphi_{n_k}\| \leq 2\|\varphi_{n_k} - \varphi\| \to 0$, which in turn implies that $\varphi = 0$, a contradiction. □

The converse of Corollary 5.7.11 is not true in general. A typical example is given by $\ell^1(\mathbb{N})$, separable, with dual $\ell^\infty(\mathbb{N})$, non-separable (Exercises 5.6.6 and 2.8.8).

Now we move into considering the Hahn–Banach Theorem in locally convex spaces. It is not immediately clear what can be done, though at least Theorem 5.7.5 is phrased in terms of a seminorm.

We will see later that surjective, continuous (i.e. bounded) operators on Banach spaces are open maps (Theorem 6.3.5). The case of linear functionals admits an easier proof and more relaxed hypotheses, and it will be of use right away.

**5.7.12. Proposition.**

*Let $\mathcal{X}$ be a topological vector space and $\varphi : \mathcal{X} \to \mathbb{C}$ a nonzero linear functional (not necessarily continuous). Then $\varphi$ is an open map. The result still holds when $\varphi$ is real linear.*

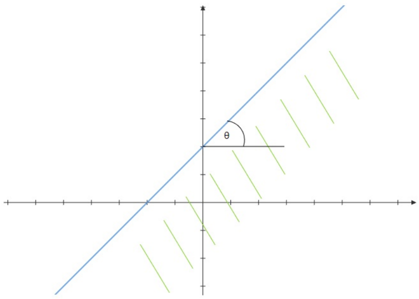
*Proof.* Since $\varphi$ is linear, it is enough to show that $0 \in \mathbb{C}$ is in the interior of the image of every open neighbourhood $Z$ of $0 \in \mathcal{X}$ (Exercise 5.7.7). From Lemma 5.4.4 we can assume that $Z$ is balanced and $Z + Z \subset Z$. Fix $z \in Z$; if $|\lambda| \leq |\varphi(z)|$, let $t \in [0, 1]$ with $|\lambda| = t\,|\varphi(z)|$. Then, letting $\alpha, \gamma$ be the arguments of $\lambda$ and $\varphi(z)$ respectively, we get $\lambda = \varphi(te^{i(-\gamma+\alpha)}z)$. As $Z$ is balanced, $te^{i(\alpha-\gamma)}z \in Z$, so $\lambda \in \varphi(Z)$. This shows that $\varphi(Z) = \bigcup_{z\in Z} \overline{B_{|\varphi(z)|}(0)}$, for the reverse inclusion is trivial. Let $\beta = \sup\{|\varphi(z)| : \ z \in Z\}$. If $\beta = \infty$, then $\varphi(Z) = \mathbb{C}$, which is open. If $\beta < \infty$ and the supremum is not achieved, then $\varphi(Z)$ is an open disk. We will prove that this supremum can indeed never be achieved. If it were, there would exist $z_0 \in Z$ with $|\varphi(z_0)| = \beta$. As $Z$ is balanced, there exists $y$ such that $\beta = \varphi(e^{iy}z_0)$, $e^{iy}z_0 \in Z$. As $Z$ is open, we can find a balanced open neighbourhood $Z_0$ of $0$ with $e^{iy}z_0 + Z_0 \subset Z$. Now let $z_1 \in Z_0$ with $\varphi(z_1) \neq 0$ (Exercise 5.7.5). Then there exists $\delta$ with $\varphi(e^{i\delta}z_1) > 0$ (as $Z_0$ is balanced), and we get

$$\beta + \varphi(e^{i\delta}z_1) = |\beta + \varphi(e^{i\delta}z_1)| = |\varphi(e^{iy}z_0 + e^{i\delta}z_1)| \leq \beta.$$

As $\varphi(e^{i\delta}z_1) > 0$, we got a contradiction.

When $\varphi$ is real linear, we can either adapt the above proof to the real case, or we can define $\tilde{\varphi}(x) = \varphi(x) - i\varphi(ix)$, which is linear by Lemma 5.7.1. Then $\varphi$ is the real part of an open map by the first part of the proof, and is thus open (Exercise 5.7.6). □

The next two results are usually quoted as “geometric” or “separation” Hahn–Banach theorems. They are the functional analysis version of the fact that we can separate two convex sets by a plane (in $\mathbb{R}^3$) or by a line (in $\mathbb{R}^2$).




Region below a line

In $\mathbb{R}^2$, consider a line $L = \{t\vec{v} + c\,e_2 : \ t \in \mathbb{R}\}$ for fixed unit vector $\vec{v} \in \mathbb{R}^2$ and $c \in \mathbb{R}$. We want to characterize the set $E$ of points below the line. At first sight this doesn't look too pretty to do analytically, but it is not that bad. We have $\vec{v} = \cos\theta\, e_1 + \sin\theta\, e_2$. If we denote by $R_\theta$ the rotation of $\theta$ radians and apply the affine transformation

$$T\vec{w} = R_{-\theta}(\vec{w} - ce_2) + ce_2$$

to $E$, the line $L$ is transformed to $L' = \mathbb{R}e_1 + ce_2$ and so $E$ is transformed to the set $E'$ of vectors below the line $L'$. These we can characterize easily, as they are those vectors such that their vertical component is at most $c$:

$$E' = \{\vec{w} : \ \langle \vec{w}, e_2\rangle \leq c\}.$$

We get that $\vec{w} \in E$ if and only if $\langle T\vec{w}, e_2\rangle \leq c$. Expanding what $T$ is, we get $\langle R_{-\theta}\vec{w}, e_2\rangle \leq c\,\langle R_{-\theta}e_2, e2\rangle = c\,\cos\theta$. Hence we can write, with $v_0 = R_\theta e_2$,

$$E = \{\vec{w} : \ \langle \vec{w}, v_0\rangle \leq c\,\cos\theta\}.$$

We learned in Section 4.5 that any linear functional on a finite-dimensional Hilbert space is given by inner product against a fixed vector. Thus we have shown that $V$ can be described as

$$E = \{\vec{w} : \ \varphi(\vec{w}) \leq c\}$$

for an appropriate $\varphi \in (\mathbb{R}^2)^*$ and $c \in \mathbb{R}$. Because of the above reasoning and the fact that it can be repeated for $\mathbb{R}^3$ with a plane instead of a line, given a vector space $V$ and a linear functional $\phi \in V^*$, a set of the form $\{x : \ \operatorname{Re}\varphi(x) = c\}$ is called a **hyperplane**. It is intuitively true, and not too hard to prove, that in $\mathbb{R}^n$ any two disjoint convex sets can be separated by a hyperplane. The result holds in way more generality, though:

***5.7.13. Theorem. (Geometric/Separation Hahn–Banach, v1)***

*Let $\mathcal{X}$ be a topological vector space, and $V, W$ disjoint, nonempty convex subsets of $\mathcal{X}$, with $V$ open. Then there exist $\varphi \in \mathcal{X}^*$ and $c \in \mathbb{R}$ such that*

$$\operatorname{Re}\varphi(v) < c \leq \operatorname{Re}\varphi(w), \quad v \in V, w \in W.$$

*Proof.* Fix $v_0 \in V$, $w_0 \in W$. Let $x_0 = w_0 - v_0$, and let $Z_0 = V - W + x_0$, $Z = Z_0 \cap (-Z_0)$. The key observation is that $Z$ is a convex open neighbourhood of 0 (Exercise 5.7.8). It is then automatic that $Z$ is absorbing. So we can consider the Minkowski functional $\mu_Z$. Combining Proposition 5.4.9 with the fact that $Z = -Z$ and thus $\mu_Z(-x) = \mu_Z(x)$ for all $x \in \mathcal{X}$, we have that $\mu_Z$ is a real seminorm. Also, $Z = \{x \in \mathcal{X} : \mu_Z(x) < 1\}$. As $V \cap W = \varnothing$, $x_0 \notin Z$ and so $\mu_Z(x_0) \geq 1$.

Let $M = \mathbb{R}x_0$, and define $\varphi_0(ax_0) = a$ on $M$. We have

$$|\varphi_0(ax_0)| = |a| \leq |a|\,\mu_Z(x_0) = \mu_Z(ax_0).$$

Then $\varphi_0$ is a real linear functional on $M$, bounded by the real seminorm $\mu_Z$. By the real part of Hahn–Banach (Theorem 5.7.5), there exists $\varphi_1 : \mathcal{X} \to \mathbb{R}$, linear, with $\varphi_1|_M = \varphi_0$ and $|\varphi_1(x)| \leq \mu_Z(x)$ for all $x \in \mathcal{X}$. Since $|\varphi_1| \leq 1$ on an open neighbourhood of 0, $\varphi$ is continuous (Exercise 5.5.8).

Now, for any $v \in V$, $w \in W$, using that $\varphi(x_0) = 1$ we get

$$\begin{aligned} \varphi_1(v) - \varphi_1(w) + 1 = \varphi_1(v - w + x_0) &\leq |\varphi_1(v - w + x_0)| \\ &\leq \mu_Z(v - w + x_0) < 1, \end{aligned}$$

since $v - w + x_0 \in Z$. This shows that $\varphi_1(v) < \varphi_1(w)$. In particular, $\varphi_1(V)$ is bounded above; let $c = \sup \varphi_1(V)$. We have that $\varphi_1(V)$ is open in $\mathbb{R}$ by Proposition 5.7.12; this implies that $c \notin \varphi_1(V)$. Thus

$$\varphi_1(v) < c \leq \varphi_1(w), \qquad v \in V, w \in W. \tag{5.24}$$

Define a linear map $\varphi : \mathcal{X} \to \mathbb{C}$ by

$$\varphi(x) = \varphi_1(x) - i\varphi_1(ix).$$

This map is continuous since $\varphi_1$ is continuous, and by Lemma 5.7.1 we have that it is linear, so $\varphi \in \mathcal{X}^*$. The inequalities (5.24) are now

$$\operatorname{Re}\varphi(v) < c \leq \operatorname{Re}\varphi(w), \qquad v \in V, w \in W. \qquad \square$$

*5.7.14. Remark* . Neither the constant $c$ nor which side is left or right in the inequalities in Theorem 5.7.13 are relevant, as things can be absorbed by changing the functional as long as $c \neq 0$. Indeed, if

$$\operatorname{Re}\varphi(v) < c \leq \operatorname{Re}\varphi(w), \qquad v \in V, w \in W,$$

replacing $\varphi$ with $\psi = (c/d)\varphi$ for nonzero $d$ such that $d/c > 0$ we get

$$\operatorname{Re}\psi(v) < d \leq \operatorname{Re}\psi(w), \qquad v \in V, w \in W.$$

And we can also revert the inequalities by replacing $c$ with $-c$ and $\varphi$ with $-\varphi$.

*5.7.15. Remark.* We cannot expect a stricter separation in Theorem 5.7.13, in the sense of having $\operatorname{Re}\varphi(v) < c < \operatorname{Re}\varphi(w)$ for all $v \in V$, $w \in W$. For instance let $\mathcal{X} = \mathbb{C}$, $V = \{z : \operatorname{Re} z > 0\}$, open and convex, and $W = \{z \in \operatorname{Re} z \leq 0\}$. Note that linear functionals, in this case, are precisely multiplication by a fixed scalar. As $0 \in W$, the strict inequality would require $\operatorname{Re}\lambda z < c < 0$ for all $z$ such that $\operatorname{Re} z > 0$, which is impossible.

*5.7.16. Remark*. Theorem 5.7.13 can fail if none of the sets is open, even if one of them is compact. For instance let $\mathcal{X} = c_0$, $W = \{0\}$, and

$$V = \{f \in c_{00} : \ f(k) > 0, \ \text{ where } k = \max\{j : \ f(j) \neq 0\}.$$

Then $V$ is convex. Let us suppose that we have $\varphi \in \mathcal{X}^*$, $c \in \mathbb{R}$ such that $\operatorname{Re}\varphi(v) < c \leq \operatorname{Re}\varphi(w)$ for all $v \in V$, $w \in W$. As $W = \{0\}$, we have

$$\operatorname{Re}\varphi(v) < c \leq 0, \quad v \in V. \tag{5.25}$$

For any $x \in c_{00}$, let $r_x = \max\{j : \ x(j) \neq 0\}$. Fix $\varepsilon > 0$. Then $\pm x + \varepsilon\, e_{r_x+1} \in V$. So

$$\pm\operatorname{Re}\varphi(x) + \varepsilon\varphi(e_{r_x+1}) \leq 0.$$

As this can be done for any $\varepsilon > 0$, we get that $\pm\operatorname{Re}\varphi(x) \leq 0$, and so $\operatorname{Re}\varphi(x) = 0$. This can also be done for $ix$, so $0 = \operatorname{Re}\varphi(ix) = -\operatorname{Im}\varphi(x)$. Therefore $\varphi(x) = 0$ for all $x \in c_{00}$. As $c_{00}$ is dense in $c_0$ and $\varphi$ is bounded, $\varphi = 0$; and this shows that separation between $\{0\}$ and $V$ as in (5.25) is not possible.

Of course, $V$ above is neither open nor closed.

*5.7.17. Remark.* Theorem 5.7.13 does not detract from the fact that an arbitrary topological vector space can be fairly pathological. We saw in Proposition 5.6.17 that if $\mathcal{X} = L^p[0,1]$ with $0 < p < 1$, then $\mathcal{X}^* = \{0\}$. Thus Theorem 5.7.13 cannot possibly work here! There is no contradiction, though, since $\mathcal{X}$ has no nontrivial convex open sets (Lemma 5.6.16).

For better topological vector spaces, Theorem 5.7.13 can be improved:

**5.7.18. Theorem. (Geometric/Separation Hahn–Banach, v2)**

*Let $\mathcal{X}$ be a locally convex space, and $A, B$ disjoint, nonempty convex subsets of $\mathcal{X}$, with $A$ compact and $B$ closed. Then there exists $\varphi \in \mathcal{X}^*$ and $c, d \in \mathbb{R}$ with*

$$\operatorname{Re}\varphi(v) < c < d < \operatorname{Re}\varphi(w), \quad v \in A, w \in B.$$

*Proof.* By Theorem 5.4.6 and the fact that the topology is locally convex, there exists a convex neighbourhood of 0 with $(A+V) \cap (B+V) = \varnothing$. By Theorem 5.7.13, there exists $\varphi \in \mathcal{X}^*$ and $d' \in \mathbb{R}$ such that $\operatorname{Re}\varphi(v) < d' \leq \operatorname{Re}\varphi(w)$ for all $v \in A+V$, $w \in B+V$. By Lemma 5.4.4 we may assume that $V$ is balanced. Since $\varphi \neq 0$, there exists $v_0 \in V$ with $\varphi(v_0) > 0$ (because $V$ is absorbing, $\varphi$ cannot be zero everywhere in $V$; for $v_1 \in V$ with $\varphi(v_1) \neq 0$, write $|\varphi(v_1)| = e^{i\gamma}\varphi(v_1)$ and define $v_0 = e^{i\gamma}v_1$). For any $a \in A$, $b \in B$,

$$\operatorname{Re}\varphi(a) < \operatorname{Re}\varphi(a) + \operatorname{Re}\varphi(v_0) = \operatorname{Re}\varphi(a+v_0) < d' \leq \operatorname{Re}\varphi(b).$$

Now, choosing $c = d' - \operatorname{Re}\varphi(v_0)$, and any $d$ with $c < d < d'$, we are done. ☐

The separation theorems are used often. Here is one of the most common forms in which they are used—the reader who perseveres through the text will see this result used again and again.

**5.7.19. Corollary.**

*If $\mathcal{X}$ is a locally convex space, $\mathcal{Y}$ is a non-dense subspace of $\mathcal{X}$, and $x \in \mathcal{X} \setminus \overline{\mathcal{Y}}$, then there exists $\varphi \in \mathcal{X}^*$ with $\varphi(x) = 1$, $\varphi|_{\overline{\mathcal{Y}}} = 0$.*

*Proof.* Applying Theorem 5.7.18 to $A = \{x\}$, $B = \overline{\mathcal{Y}}$, we get $\varphi' \in \mathcal{X}^*$ with $\operatorname{Re}\varphi'(x) \notin \operatorname{Re}\varphi'(\mathcal{Y})$. Since $\mathcal{Y}$ is a subspace, this implies that $\operatorname{Re}\varphi'(x) \neq 0$, and it also implies that $\operatorname{Re}\varphi'(\mathcal{Y})$ is a proper subspace of $\mathbb{R}$; hence $\operatorname{Re}\varphi'(\mathcal{Y}) = \{0\}$. As $\mathcal{Y}$ is a complex vector space, we get that $\varphi'(\mathcal{Y}) = \{0\}$ (using Lemma 5.7.1). Now put $\varphi = (1/\varphi'(x))\varphi'$. ☐

*5.7.20. Remark.* An alternative argument to prove Corollary 5.7.19 is to use Theorem 5.7.18 to produce $\varphi \in \mathcal{X}^*$ and $c \in \mathbb{R}$ with $\operatorname{Re}\varphi(y) < c < \operatorname{Re}\varphi(x)$ for all $y \in \mathcal{Y}$. As the inequality applies both to $y$ and $-y$, we have $-c < \operatorname{Re}\varphi(y) < c$. But $ny \in \mathcal{Y}$ for arbitrarily large $n$, showing that $\operatorname{Re}\varphi(y) = 0$. Then $c > 0$ and $\operatorname{Re}\varphi(x) \neq 0$, so $\varphi(x) \neq 0$.

*5.7.21. Remark.* When $\mathcal{X}$ is normed, Corollary 5.7.19 is a direct consequence of Hahn–Banach (Theorem 5.7.5), via Lemma 5.7.10. But the problem with applying Theorem 5.7.5 here is that using some seminorm—there are many available, as the topology is determined by a whole family of them—we cannot guarantee the continuity of the extension.

*5.7.22. Example.* Corollary 5.7.19 can fail when the topological vector space is not locally convex. For instance let $\mathcal{X} = L^p[0,1]$ with $0 < p < 1$. Let $\mathcal{Y} \subset \mathcal{X}$ be any proper subspace and $x \in \mathcal{X} \setminus \mathcal{Y}$. As $\mathcal{X}^* = \{0\}$ (Proposition 5.6.17), no functional with $\varphi(x) = 1$ can exist.

Next comes a straightforward generalization of Lemma 5.7.8

***5.7.23. Lemma.***

*Let $\mathcal{X}$ be an infinite-dimensional locally convex space. Then $\mathcal{X}^*$ is infinite-dimensional.*

*Proof.* The proof is exactly that of Lemma 5.7.8, with Corollary 5.7.19 used to produce the $\varphi_n$. □

Another consequence of the geometric Hahn–Banach theorems is the following corollary, which shows that we can still extend continuous functionals in locally convex spaces as in the Hahn–Banach Theorem (no control on the norm, though, as there is no norm in general).

***5.7.24. Corollary.***

*Let $\mathcal{X}$ be a locally convex space, $\mathcal{Y}$ a (not necessarily closed) subspace of $\mathcal{X}$, $\varphi \in \mathcal{Y}^*$. Then there exists $\tilde{\varphi} \in \mathcal{X}^*$ with $\tilde{\varphi}|_{\mathcal{Y}} = \varphi$.*

*Proof.* If $\varphi = 0$, then take $\tilde{\varphi} = 0$. Otherwise, let $\mathcal{Y}_0 = \ker \varphi \subset \mathcal{Y}$. Note that $\mathcal{Y}_0$ is closed in $\mathcal{Y}$, because $\varphi \in \mathcal{Y}^*$. Since $\varphi \neq 0$, there exists $y_0 \in \mathcal{Y}$ with $\varphi(y_0) = 1$. The fact that $\varphi$ is continuous implies that $y_0 \notin \overline{\mathcal{Y}_0}$ in $\mathcal{X}$: this requires a tiny bit of care because $\mathcal{Y}$ is not necessarily closed, so it is not necessarily true that closed subsets in the relative topology of $\mathcal{Y}$ are closed in $\mathcal{X}$. But there is no issue in this case: if $y_0 \in \overline{\mathcal{Y}_0}$ (the closure in $\mathcal{X}$), then $y_0 \in \overline{\mathcal{Y}_0} \cap \mathcal{Y}$ which is the closure of $\mathcal{Y}_0$ in $\mathcal{Y}$; then the continuity of $\varphi$ in $\mathcal{Y}$ would imply that $\varphi(y_0) = 0$, a contradiction.

By Corollary 5.7.19 there exists $\tilde{\varphi} \in \mathcal{X}^*$ with $\tilde{\varphi}(y_0) = 1$, $\tilde{\varphi}|_{\mathcal{Y}_0} = 0$. Then, for any $y \in \mathcal{Y}$, using that $y - \varphi(y)y_0 \in \mathcal{Y}_0$,

$$\tilde{\varphi}(y) = \tilde{\varphi}(y - \varphi(y)y_0) + \tilde{\varphi}(\varphi(y)y_0) = \varphi(y)\tilde{\varphi}(y_0) = \varphi(y). \qquad \square$$

When $\mathcal{Y} \subset \mathcal{X}$ is dense, we do not need local convexity in Corollary 5.7.24:

***5.7.25. Proposition.***

*Let $\mathcal{X}$ be a topological vector space and $\mathcal{X}_0 \subset \mathcal{X}$ a dense subspace. Then $\mathcal{X}_0^* = \mathcal{X}^*$.*

*Proof.* We need to show that $\varphi \in \mathcal{X}_0^*$ can be extended to $\mathcal{X}$ and still continuous. Fix $x \in \mathcal{X} \setminus \mathcal{X}_0$. Let $\{x_j\} \subset \mathcal{X}_0$ be a net with $x_j \to x$. Given $\varepsilon > 0$, there exists an open neighbourhood $V$ of 0 in $\mathcal{X}$ such that $v \in \mathcal{X}_0 \cap V$ implies $|\varphi(v)| < \varepsilon$ (the open sets of $\mathcal{X}_0$ are of the form $\mathcal{X}_0 \cap V$ for $V$ open in $\mathcal{X}$). By Lemma 5.4.4 there exists $W \subset V$ balanced and such that $W + W \subset V$. Because $x_j \to x$, there exists $j_0$ such that $x_j - x \in W$ for all $j \geq j_0$ and so $x_j - x_k \in V$ for all $j, k \geq j_0$. Then $|\varphi(x_j) - \varphi(x_k)| = |\varphi(x_j - x_k)| < \varepsilon$ and so the net of numbers $\{\varphi(x_j)\}$ is Cauchy in $\mathbb{C}$. Let $\tilde{\varphi}(x) = \lim_j \varphi(x_j)$. If $\{x'_\ell\} \subset \mathcal{X}_0$ is another net that converges to $x$, there exists an $\ell_0$ such that $x'_\ell - x \in W$ for all $\ell \geq \ell_0$. Then, for $j \geq j_0$ and $\ell \geq \ell_0$ we have $x_j - x'_\ell = (x_j - x) + (x - x'_\ell) \in W + W \subset V$, so

$$|\varphi(x_j) - \varphi(x'_\ell)| = |\varphi(x_j - x'_\ell)| < \varepsilon.$$

As this can be done for any $\varepsilon$, this shows that $\tilde{\varphi}(x)$ is well-defined, in the sense that the limit is the same for any net that converges to $x$. This uniqueness also shows that $\tilde{\varphi}$ is linear. As for continuity, given $\varepsilon > 0$ and $V$ as before, use Lemma 5.4.4 twice to get $W \subset V$, open, balanced, with $0 \in W$ and $W + W + W + W \subset V$. If $x - y \in W$, we may fix nets $\{x_j\}, \{y_k\} \subset \mathcal{X}_0$ that converge to $x$ and $y$ respectively, and so there exist $j_0$ and $k_0$ such that $x - x_j \in W$, $|\tilde{\varphi}(x) - \varphi(x_j)| < \varepsilon$ and $y - y_k \in W$, $|\tilde{\varphi}(y) - \varphi(y_k)| < \varepsilon$ if $j \geq j_0$, $k \geq k_0$. Then

$$x_j - y_k = (x_j - x) + (x - y) + (y - y_k) + 0 \in W + W + W + W \subset V,$$

and so

$$|\tilde{\varphi}(x) - \tilde{\varphi}(y)| \leq |\tilde{\varphi}(x) - \varphi(x_j)| + |\varphi(x_j - y_k)| + |\varphi(y_k) - \tilde{\varphi}(y)| < 3\varepsilon. \qquad \square$$

For of Banach spaces, Proposition 5.7.25 was proven in Exercise 5.5.11. The argument is no different in a topological vector space.

*5.7.26. Remark* . When $\mathcal{X}$ is not locally convex, Proposition 5.7.25 has an interesting consequence. Fix $p \in (0, 1)$ and consider $\mathcal{X} = L^p[0, 1]$. Then $\mathcal{X}$ is a (metric!) topological vector space, as shown on subsection 5.6.7. We have that $\mathcal{X}^* = \{0\}$ by Proposition 5.6.17. Now let $\mathcal{X}_0 = C[0,1]$. By Proposition 5.6.18, $\mathcal{X}_0$ is dense $\mathcal{X}$. Then Proposition 5.7.25 tells us that no nonzero linear functional $\varphi : C[0,1] \to \mathbb{C}$ can be continuous for $d_p$. Things as nice as

$$\phi : f \longmapsto \int_0^1 f$$

are discontinuous for the metric $d_p$ (cfr. Exercise 5.7.12). Any other way that we may come up with to assign numbers to continuous functions in a linear way, will be discontinuous.

The following consequence of Hahn–Banach will prove useful down the road.

**5.7.27. Proposition.**

*Let $\mathcal{X}$ be a normed space.*

*(i) If $M \subset \mathcal{X}$ is a finite dimensional space, then $M$ is topologically complemented.*

*(ii) If $V \subset \mathcal{X}$ is a closed subspace with $m = \dim \mathcal{X}/V < \infty$, then $V$ is topologically complemented with the complement of dimension $m$.*

*Proof.* Exercises 5.7.2 and 5.7.3. ☐

### 5.7.1. Exercises.

(5.7.1) Write a complete proof of Corollary 5.7.6.

(5.7.2) Let $\mathcal{X}$ be a locally convex space and $\mathcal{Y} \subset \mathcal{X}$ a subspace with $\dim \mathcal{Y} < \infty$. Show that $\mathcal{Y}$ is (topologically) complemented.

(5.7.3) Let $\mathcal{X}$ be a locally convex space and $V \subset \mathcal{X}$ a closed subspace. Show that if $m = \dim \mathcal{X}/V < \infty$, then $V$ is topologically complemented with complement of dimension $m$. Give an example to show that $V$ need not be complemented if it is not closed.

(5.7.4) Show that $\ell^\infty(\mathbb{N})^* \neq \ell^1(\mathbb{N})$, by proving that the equality would imply that $\ell^\infty(\mathbb{N})/c_0$ has trivial dual, in contradiction with Corollary 5.7.7.

(5.7.5) Let $\mathcal{X}$ be a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ linear. Show that if $\varphi|_V = 0$ on some neighbourhood $V$ of 0, then $\varphi = 0$.

(5.7.6) *(This is part of the proof of Proposition 5.7.12)* Show that if $\mathcal{X}$ is a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ is linear and open, then $\operatorname{Re}\varphi : \mathcal{X} \to \mathbb{R}$ is real linear and open.

(5.7.7) *(This is part of the proof of Proposition 5.7.12)* Prove that if $\mathcal{X}$ is a TVS and $\varphi : \mathcal{X} \to \mathbb{C}$ is a linear functional, then $\varphi$ is an open map if and only if for every neighbourhood $Z$ of $0 \in \mathcal{X}$, $\varphi(Z)$ contains $0 \in \mathbb{C}$ as an interior point.

(5.7.8) Prove that the set $Z$ used at the beginning of the proof of Theorem 5.7.13 is nonempty, open, and convex.

(5.7.9) Is the condition "$V$ open" necessary for Theorem 5.7.13?

(5.7.10) Show that if $\mathcal{X}$ is a locally convex space and $M \subset \mathcal{X}$ is a subspace, then the closure $\overline{M}$ of $M$ is

$$\overline{M} = \bigcap_{f \in K_M} \ker f,$$

where $K_M = \{f \in \mathcal{X}^* : \; M \subset \ker f\}$.

(5.7.11) Use Exercise 5.7.10 to show that if $\mathcal{X}$ is locally convex and $M \subset \mathcal{X}$ is a subspace, then $M$ is dense in $\mathcal{X}$ if and only $\{f \in \mathcal{X}^* : \; f = 0 \text{ on } M\} = \{0\}$.

(5.7.12) In Remark 5.7.26 it is shown that the map $f \longmapsto \int_0^1 f$ is not continuous on $\mathcal{X} = L^p[0,1]$, $0 < p < 1$. Prove this explicitly by finding a sequence $\{f_n\} \subset \mathcal{X}$ such that $d_p(f_n, 0) \to 0$ while $\int_0^1 |f_n| \nearrow \infty$.

(5.7.13) Let $\mathcal{X}$ be a locally convex space, and $x, y \in \mathcal{X}$. Show that if $\operatorname{Re}\varphi(x) = \operatorname{Re}\varphi(y)$ for all $\varphi \in \mathcal{X}^*$, then $x = y$.

(5.7.14) Let $\mathcal{Y}$ be a balanced, convex, closed subset of a locally convex space $\mathcal{X}$, and let $x \in \mathcal{X} \setminus \mathcal{Y}$. Show that there exists $\varphi \in \mathcal{X}^*$ with $\varphi(x) > 1$ and $|\varphi(y)| \le 1$ for all $y \in \mathcal{Y}$.

CHAPTER 6

# Huge consequences of Baire's Category Theorem

Baire's Category Theorem (6.3.1) is a neat result about complete metric spaces—or even just locally compact spaces, cfr. Exercise 6.3.2—that is the basis to prove many fundamental results in functional analysis.

Since most of the consequences we use are about bounded linear operators, we begin with a brief panorama on them. A deeper look at bounded linear operators will have to wait until Chapters 9 and 10.

## 6.1. Bounded linear operators

Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. We will denote by $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ the set of all continuous linear functions $\mathcal{X} \to \mathcal{Y}$. When $\mathcal{Y} = \mathcal{X}$, we write $\mathcal{B}(\mathcal{X})$. We have already encountered $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{C})$.

A common convention is to write the action of an element of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ acting on an element of $\mathcal{X}$ without brackets. So, if $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $x \in \mathcal{X}$, we write $Tx$ instead of $T(x)$. Convention is less standard regarding the use of uppercase or lowercase letters for the operators, but we will stay with uppercase for now.

**6.1.1. Definition.**

*Given $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, we denote the kernel and the range of $T$ by*

$$\ker T = \{x \in \mathcal{X} : \ Tx = 0\}, \qquad \operatorname{ran} T = \{Tx : \ x \in \mathcal{X}\}$$

*respectively.*

The big results in this chapter are concerned with invertibility in $\mathcal{B}(\mathcal{X})$. Before moving into deeper stuff, we start with a few elementary results about operators. We saw in Proposition 4.5.1 (and the preceding discussion on page 331) and Proposition 5.5.4 that in infinite dimension linearity does not imply automatic continuity. The functionals considered there are in $\mathcal{B}(\mathcal{X}, \mathbb{C})$, and the ideas from those results extend to $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

***6.1.2. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be normed vector spaces, and let $T : \mathcal{X} \to \mathcal{Y}$ be linear. Then the following statements are equivalent:*

*(i) $T$ is continuous;*

*(ii) $T$ is continuous at $0$;*

*(iii) there exists $x \in \mathcal{X}$ such that $T$ is continuous at $x$;*

*(iv) there exists $r > 0$ such that, for every $x \in \mathcal{X}$, $\|Tx\| \leq r\,\|x\|$.*

*Proof.* Exercise 6.1.2. □

In analogy with the case of functionals, we define

**6.1.3. Definition.**

*A linear operator $T : \mathcal{X} \to \mathcal{Y}$ that satisfies any of the statements in Proposition 6.1.2 is called a* ***bounded operator****, and the set of such operators is denoted by $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

*Given $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, the* ***norm*** *of $T$ is the number*

$$\|T\| = \inf\{r : \|Tx\| \leq r\|x\| \text{ for all } x \in \mathcal{X}\}. \tag{6.1}$$

*When $\|T\| \leq 1$ we say that $T$ is* ***contractive****.*

Of course one needs to prove that such assignment is indeed a norm (Exercise 6.1.4). And again we can mimic the characterizations from Proposition 5.5.6:

***6.1.4. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then*

*(i)* $\|T\| = \min\{r : \|Tx\| \leq r\|x\| \text{ for all } x \in \mathcal{X}\}$;

*(ii)* $\|T\| = \sup\{\|Tx\| : \ \|x\| = 1\}$;

*(iii)* $\|T\| = \sup\{\|Tx\|/\|x\| : \ x \neq 0\}$;

*(iv) for any* $x \in \mathcal{X}$, $\|Tx\| \leq \|T\|\,\|x\|$;

*(v) if* $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$*, Then* $\|ST\| \leq \|S\|\,\|T\|$.

*Proof.* Exercise 6.1.5. □

When $\mathcal{Y}$ is "good enough", $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a Banach space. Indeed, the next result has Proposition 5.5.8 as a particular case.

***6.1.5. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces with $\mathcal{Y}$ Banach. Then $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a Banach space with the norm defined in* (6.1).

*Proof.* Exercise 6.1.6. □

*6.1.6. Remark.* It is not clear a priori how big or small can $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ be. When $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional, we can identify $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ with the algebra $M_{m,n}(\mathbb{C})$ of $m \times n$ matrices, where $n = \dim \mathcal{X}$, $m = \dim \mathcal{Y}$ (Exercise 6.1.7). In the general case of course it depends a lot on the space $\mathcal{X}$, but $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ can never be trivial. For starters, given any $\varphi \in \mathcal{X}^*$, $y_0 \in \mathcal{Y}$ (recall that $\mathcal{X}^*$ is always fairly big, by Corollary 5.7.7), the map $T : x \mapsto \varphi(x)y_0$ is in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$; one can even show that $\|T\| = \|\varphi\|\,\|y_0\|$ (Exercise 6.1.10). More generally, given $\varphi_1, \ldots, \varphi_n \in \mathcal{X}^*$ and $y_1, \ldots, y_n \in \mathcal{Y}$, the map $x \mapsto \sum_j \varphi_j(x)y_j$ is in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$. Even more generally, if $\mathcal{Y}$ is Banach and $\varphi_1, \varphi_2, \ldots \in \mathcal{X}^*$ and $y_1, y_2, \ldots \in \mathcal{Y}$ with $\sum_j \|\varphi_j\|\,\|y_j\| < \infty$, then the map $x \mapsto \sum_j \varphi_j(x)y_j$ is in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ (Exercise 6.1.10). All these are examples of **compact operators**. These and other compact operators will be discussed further down the road (cfr. Sections 9.6 and 10.6).

When $\mathcal{Y} = \mathcal{X}$, there is an additional canonical element in $\mathcal{B}(\mathcal{X})$, namely the **identity map** $\mathrm{id}_{\mathcal{X}}$ given by $\mathrm{id}_{\mathcal{X}} x = x$ for all $x \in \mathcal{X}$. We will often write $I$ instead of $\mathrm{id}_{\mathcal{X}}$, to emphasize the fact that $I$ is the identity element when $\mathcal{B}(\mathcal{X})$ is seen as an algebra.

There exist Banach spaces $\mathcal{X}$ such that $\mathcal{B}(\mathcal{X})$ consists of nothing more than sums of compact operators and scalar multiples of the identity [AH11], but for the most common Banach spaces $\mathcal{B}(\mathcal{X})$ is much bigger than that.

*6.1.7. Example*. Let us mention a couple of examples of operators which are not of the form described in Remark 6.1.6, and yet they are familiar. Let $\mathcal{X} = C[0,1]$, with the supremum norm, and $T : \mathcal{X} \to \mathcal{X}$ the operator given by $(Tf)(x) = \int_0^x f$. We have

$$\left| \int_0^x f \right| \leq \int_0^x |f| \leq x \, \|f\|_\infty \leq \|f\|_\infty.$$

So $\|Tf\| \leq \|f\|_\infty$, which implies that $\|T\| \leq 1$ (actually, $\|T\| = 1$, as seen by using $f = 1$). A second familiar example comes from taking $\mathcal{X} = \mathcal{Y} = \mathbb{C}[x]$, with the norm $\|p\|_{\mathcal{X}} = \|p\|_\infty + \|p'\|_\infty$ on $\mathcal{X}$ and the norm $\|p\|_{\mathcal{Y}} = \|p\|_\infty$ on $\mathcal{Y}$, and letting $D : \mathcal{X} \to \mathcal{X}$ be given by $Dp = p'$. Then

$$\|Dp\|_{\mathcal{Y}} = \|p'\|_\infty \leq \|p\|_\infty + \|p'\|_\infty = \|p\|_{\mathcal{X}}.$$

Hence $D$ is bounded with $\|D\| \leq 1$.

*6.1.8. Remark*. The space $\mathcal{B}(\mathcal{X})$ becomes an algebra over $\mathbb{C}$ when the product $TS$ is to be taken as the composition of functions, i.e. $TSx$ is $T(Sx)$. This means that $\mathcal{B}(\mathcal{X})$ is a ring with identity that is also a complex vector space that moreover satisfies $(\alpha T)(\beta S) = (\alpha\beta)TS$ for all $\alpha, \beta \in \mathbb{C}$, $T, S \in \mathcal{B}(X)$. This is often summarized by saying that $\mathcal{B}(\mathcal{X})$ is an **algebra**.

In this setting a consequence of the definition of the norm in $\mathcal{B}(\mathcal{X})$, as seen in Proposition 6.1.4, is that $\|ST\| \leq \|S\| \, \|T\|$ for all $S, T \in \mathcal{B}(\mathcal{X})$. That is, the norm is submultiplicative.

The following extension result (that generalizes Exercise 5.5.11) is used frequently and without mention when dealing with bounded linear operators.

**6.1.9. Proposition.**

*Let $\mathcal{X}$ be a normed space, $\mathcal{Y}$ a Banach space, and $\mathcal{X}_0 \subset \mathcal{X}$ a dense subspace. Let $T : \mathcal{X}_0 \to \mathcal{Y}$ be linear and bounded. Then $T$ admits a unique extension $\tilde{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. The extension satisfies $\|\tilde{T}\| = \|T\|$.*

*Proof.* Let $x \in \mathcal{X}$ and choose $\{x_n\} \subset \mathcal{X}_0$ with $x_n \to x$. We have

$$\|Tx_n - Tx_m\| = \|T(x_n - x_m)\| \leq \|T\| \, \|x_n - x_m\|.$$

As $\{x_n\}$ is Cauchy, it follows that $\{Tx_n\}$ is Cauchy. Since $\mathcal{Y}$ is complete, there exists $y \in \mathcal{Y}$ with $Tx_n \to y$. And this $y$ is unique: if $x'_n \to x$, then

$$\|Tx'_n - y\| \leq \|Tx'_n - Tx_n\| + \|Tx_n - y\| \leq \|T\| \, \|x'_n - x_n\| + \|Tx_n - y\|$$
$$\leq \|T\| \left(\|x'_n - x\| + \|x - x_n\|\right) + \|Tx_n - y\| \xrightarrow[n\to\infty]{} 0.$$

This allows us to denote this $y$ by $\tilde{T}x$. The uniqueness also makes $\tilde{T}$ linear; indeed, if $x_n \to x_1$ and $x'_n \to x_2$, then

$$\tilde{T}(cx_1 + x_2) = \lim_n T(cx_n + x'_n) = \lim_n cTx_n + Tx'_n = c\tilde{T}x_1 + \tilde{T}x_2.$$

When $x \in \mathcal{X}_0$, the uniqueness allows us to choose a sequence $\{x_n\}$ that suits us; let us take $x_n = x$ for all $n$. Then $\tilde{T}x = \lim Tx = Tx$. Thus $\tilde{T}|_{\mathcal{X}_0} = T$. Finally, $\|\tilde{T}\| = \|T\|$, since $\|T\| \leq \|\tilde{T}\|$ (from $\mathcal{X}_0 \subset \mathcal{X}$) and

$$\|\tilde{T}x\| = \|\lim_n Tx_n\| = \lim_n \|Tx_n\| \leq \|T\| \, \lim_n \|x_n\| = \|T\| \, \|x\|. \qquad \square$$

There is a small subtlety in the proof above, which is that we used that $\|x_n\| \to \|x\|$ if $x_n \to x$. This is a straightforward consequence of the reverse triangle inequality (cfr. page 336).

We include here a result about $\ell^1(\mathbb{N})$ that we will be of use much later—it is not needed to follow the flow of this chapter. It is amazing that anything we can possibly know about any separable Banach space is somehow encoded inside $\ell^1(\mathbb{N})$. Indeed,

***6.1.10. Theorem.***

*Let $\mathcal{X}$ be a separable Banach space. Then there exists a closed subspace $\mathcal{Y} \subset \ell^1(\mathbb{N})$ such that $\mathcal{X}$ is isometrically isomorphic to $\ell^1(\mathbb{N})/\mathcal{Y}$.*

*Proof.* Let $\{x_n\}$ be a dense subset of the unit ball of $\mathcal{X}$. We will have to choose this sequence with care, but that will come a bit later. Define $T : \ell^1(\mathbb{N}) \to \mathcal{X}$ by $Ta = \sum_k a_k x_k$. This gives a bounded linear operator, since

$$\|Ta\| = \Big\| \sum_k a_k x_k \Big\| \leq \sum_k |a_k| \, \|x_k\| \leq \|a\|_1.$$

So $\|T\| \leq 1$. We claim that $T$ is surjective. To see this, let $r = 3/2$. Given any $x \in \mathcal{X}$ with $\|x\| \leq 1$, the distance from $rx$ to the unit ball is at most $1/2$. So there exists $x_{n_1}$ with $\|rx - x_{n_1}\| < 1$. Repeat the argument to find $x_{n_2}$ with $\|r(rx - x_{n_1}) - x_{n_2}\| < 1$. After $N$ steps, we have

$$\|x - \sum_{k=1}^{N} r^{-k} x_{n_k}\| < r^{-N}.$$

It follows that $x = \sum_k r^{-k} x_{n_k} = Ta$, where $a = \sum_k r^{-k} e_{n_k} \in \ell^1(\mathbb{N})$. So $T$ is surjective and thus it induces an isomorphism $\tilde{T} : \ell^1(\mathbb{N})/\mathcal{Y} \to \mathcal{X}$ by $\tilde{T}(a + \mathcal{Y}) = Ta$, where $\mathcal{Y} = \ker T$. Next we want $\tilde{T}$ to be isometric. This is not true in general, but it is if we choose $\{x_n\}$ carefully.

On the easy side we have, for any $y \in \mathcal{Y}$,

$$\|\tilde{T}(a + \mathcal{Y})\| = \|Ta\| = \|T(a + y)\| \leq \|a + y\|.$$

So $\|\tilde{T}(a + \mathcal{Y})\| \leq \inf\{\|a + y\| : \ y \in \mathcal{Y}\| = \|a + \mathcal{Y}\|$. To prove the reverse inequality, is where we are going to have to be more picky in choosing our sequence $\{x_n\}$. Let $A$ be a countable dense subset in the unit ball of $\ell^1(\mathbb{N})$ and $\mathcal{X}_0 \subset \mathcal{X}$ dense in the unit ball of $\mathcal{X}$. The set

$$\mathcal{X}_1 = \left\{ \frac{\sum_{k=1}^N a_k x_k}{\left\|\sum_{k=1}^N a_k x_k\right\|} : \ N \in \mathbb{N}; x_1, \ldots, x_N \in \mathcal{X}, \ a \in A \right\}$$

is countable and contained in the closed unit ball of $\mathcal{X}$, so $\mathcal{X}_0 \cup \mathcal{X}_1$ is countable and dense, and contained in the unit ball of $\mathcal{X}$. Now form the sequence $\{x_n\}$ by repeating each element of $\mathcal{X}_0 \cup \mathcal{X}_1$ infinitely many times. Still countable and dense!

All we did above we can do with $T$ defined in terms of the new sequence $\{x_n\}$ (note that the definition of the sequence does not use $T$ in any way, so we could have started with this; we felt it was preferable to describe what was going on first, though).

Fix $a \in A$, and $\varepsilon > 0$. There exists $N_0 \in \mathbb{N}$ with $\sum_{j>N} |a_j| < \varepsilon$ for all $N > N_0$. Because of the way we constructed the sequence, there exists $m > N$ with $x_m = \frac{\sum_{k=1}^N a_k x_k}{\left\|\sum_{k=1}^N a_k x_k\right\|}$. Now let $b \in \ell^1(\mathbb{N})$ be given by

$$b = -\sum_{k=1}^N a_k e_k + \Big\|\sum_{k=1}^N a_k x_k\Big\| e_m.$$

We have

$$Tb = -\sum_{k=1}^N a_k x_k + \Big\|\sum_{k=1}^N a_k x_k\Big\| x_m = 0,$$

so $b \in \ker T$. Then

$$\begin{aligned} \|a + \mathcal{Y}\| \leq \|a + b\|_1 = \sum_j |a_j + b_j| &= \Big|a_m + \Big\|\sum_{k=1}^N a_k x_k\Big\|\Big| + \sum_{j>N, \ j\neq m} |a_j| \\ &\leq \Big\|\sum_{k=1}^N a_k x_k\Big\| + \sum_{j>N} |a_j| \leq \varepsilon + \Big\|\sum_{k=1}^N a_k x_k\Big\|. \end{aligned}$$

As we can take $N$ arbitrarily big,

$$\|a+\mathcal{Y}\| \leq \varepsilon + \limsup_{N\to\infty} \Big\| \sum_{k=1}^{N} a_k x_k \Big\| = \varepsilon + \|Ta\|.$$

As $\varepsilon$ was arbitrary, we get $\|a+\ker T\| \leq \|Ta\|$. Finally, as this holds in a dense subset of the unit ball of $\ell^1(\mathbb{N})$ and everything is continuous, the inequality $\|a+\mathcal{Y}\| \leq \|Ta\|$ holds for any $a$ in the unit ball of $\ell^1(\mathbb{N})$, and by linearity for all $a \in \ell^1(\mathbb{N})$. So $\tilde{T}$ is an isometry. □

**6.1.1. Exercises.**

(6.1.1) Prove the First Isomorphism Theorem for linear operators: given vector spaces $\mathcal{X}, \mathcal{Y}$ and $T : \mathcal{X} \to \mathcal{Y}$ linear, then

$$\mathcal{X}/\ker T \simeq \operatorname{ran} T$$

canonically.

(6.1.2) Prove Proposition 6.1.2.

(6.1.3) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, and $T : \mathcal{X} \to \mathcal{Y}$ be linear. Show that the following statements are equivalent:

(i) $T$ is unbounded;

(ii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $\|x_n\| = 1$ and $\|Tx_n\| > n$ for all $n$;

(iii) there exists a sequence $\{x_n\} \subset \mathcal{X}$ such that $x_n \to 0$ and $\|Tx_n\| = 1$ for all $n$.

(6.1.4) Prove that (6.1) defines a norm in $\mathcal{B}(\mathcal{X}, \mathcal{Y})$, and that

$$\|Tx\| \leq \|T\|\,\|x\|, \qquad T \in \mathcal{B}(\mathcal{X}, \mathcal{Y}), \quad x \in \mathcal{X}.$$

(6.1.5) Prove Proposition 6.1.4.

(6.1.6) Prove Proposition 6.1.5 (*Hint: Proposition 5.5.8 is a particular case*).

(6.1.7) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces. Show that if $\dim \mathcal{X} = n$, $\dim \mathcal{Y} = m$, then $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ can be identified with $M_{m,n}(\mathbb{C})$.

(6.1.8) Let $\mathcal{Y}$ be an infinite-dimensional normed space. Show that the space $\mathcal{B}(\mathbb{C}, \mathcal{Y})$ is infinite-dimensional.

(6.1.9) Let $\mathcal{X}, \mathcal{Y}$ be a Banach spaces and $T : \mathcal{X} \to \mathcal{Y}$ linear and isometric, that is $\|Tx\| = \|x\|$ for all $x \in \mathcal{X}$. Show that $\operatorname{ran} T$ is closed.

(6.1.10) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces.

(i) Show that if $\varphi \in \mathcal{X}^*$, $y_0 \in \mathcal{Y}$, and $T : x \longmapsto \varphi(x)y_0$, then $\|T\| = \|\varphi\|\,\|y_0\|$.

(ii) Show that if $\varphi_1, \ldots, \varphi_n \in \mathcal{X}^*$ and $y_1, \ldots, y_n \in \mathcal{Y}$, then $T : x \longmapsto \sum_j \varphi_j(x) y_j$ is bounded.

(iii) Show that if $\mathcal{Y}$ is Banach, $\varphi_1, \varphi_2, \ldots \in \mathcal{X}^*$ and $y_1, y_2, \ldots \in \mathcal{Y}$ with $\sum_j \|\varphi_j\| \, \|y_j\| < \infty$, then $T : x \longmapsto \sum_j \varphi_j(x) y_j$ is bounded.

(6.1.11) Let $T : c_{00} \to c_{00}$ be given by

$$T(x_1, x_2, \cdots) = \Big(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots\Big).$$

Show that $T$ is bounded and bijective.

(6.1.12) Let $T : c_0 \to c_0$ be given by

$$T(x_1, x_2, \cdots) = \Big(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots\Big).$$

Show that $T$ is bounded (with $\|T\| = 1$) and injective, but not surjective.

## 6.2. Invertibility in $\mathcal{B}(\mathcal{X})$

An element $T \in \mathcal{B}(\mathcal{X})$ is said to be **invertible** if it is invertible in the algebraic sense, i.e. there exists $S \in \mathcal{B}(\mathcal{X})$ such that $TS = ST = I$. Such an inverse, if it exists, is unique (Exercise 1.1.7); this warrants the notation $T^{-1}$ for it.

We denote by $GL(\mathcal{X})$ the set of all invertible operators in $\mathcal{B}(\mathcal{X})$. It is important to remark that, as opposed to the familiar case of square matrices, *both* equalities $TS = I$ and $ST = I$ are required for invertibility, and neither implies the other in general. For instance let $\mathcal{X} = \ell^2(\mathbb{N})$, and

$$Tx = (x_2, x_3, \ldots, ), \qquad Sx = (0, x_1, x_2, \ldots).$$

Since $\|Tx\| \leq \|x\|$ and $\|Sx\| = \|x\|$ for all $x$, we have $T, S \in \mathcal{B}(\mathcal{X})$. Also, $TS = I$, but $ST \neq I$ ($T$ is not injective!). So neither $T$ nor $S$ are invertible, even though $T$ admits a right inverse and $S$ a left inverse. These operators are called the left and right **unilateral shifts**, and we will encounter them often.

**6.2.1. Lemma. (Neumann Series)**

*Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ with $0 < \|T\| < 1$. Then the sequence $\{\sum_{n=0}^k T^n\}_k$ is a Cauchy sequence, and so the limit $S = \sum_{n=0}^\infty T^n$ exists in $\mathcal{B}(\mathcal{X})$. The operator $S$ is invertible, with inverse $I - T$.*

*Proof.* We have

$$\left\| \sum_{n=k+1}^{\ell} T^n \right\| \leq \sum_{n=k+1}^{\ell} \|T^n\| \leq \sum_{n=k+1}^{\ell} \|T\|^n = \frac{\|T\|^{k+1} - \|T\|^{\ell+1}}{1 - \|T\|}.$$

As $\|T\| < 1$, the right-hand-side above can be made arbitrarily small, and so the sequence $\{\sum_{n=0}^{k} T^k\}_k$ is Cauchy in $\mathcal{B}(\mathcal{H})$ and it converges to an operator $S \in \mathcal{B}(\mathcal{H})$ by the completeness of $\mathcal{B}(\mathcal{H})$ (Proposition 6.1.5). We have

$$(I - T) \sum_{n=0}^{k} T^n = I - T^{k+1}.$$

As $\|T^{k+1}\| \leq \|T\|^{k+1}$ (Proposition 6.1.4), $T^{k+1} \to 0$ as $n \to \infty$. So $(I-T)S = I$. Similarly, $S(I - T) = I$, so $S$ is invertible with inverse $I - T$. □

*6.2.2. Remark* . The typical use of Lemma 6.2.1 is that when $\|I - T\| < 1$, then $T$ is invertible with inverse $\sum_{k=0}^{\infty}(I - T)^k$.

***6.2.3. Proposition.***

*In a Banach space $\mathcal{X}$, the set $GL(\mathcal{X})$ is open in $\mathcal{B}(\mathcal{X})$. More concretely, if $T \in GL(\mathcal{X})$ and $\|S - T\| < \|T^{-1}\|^{-1}$, then $S \in GL(\mathcal{X})$.*

*Proof.* Let $T \in GL(\mathcal{X})$. Suppose that $\|S - T\| < \|T^{-1}\|^{-1}$. Then

$$\|T^{-1}S - I\| = \|T^{-1}(S - T)\| \leq \|T^{-1}\| \, \|S - T\| < \|T^{-1}\| \, \|T^{-1}\|^{-1} = 1.$$

By Lemma 6.2.1, $T^{-1}S$ is invertible and so $S$ is invertible (Exercise 6.2.1). We have shown that all the operators in the ball $B_{\|T^{-1}\|^{-1}}(T)$ are invertible, so $GL(\mathcal{X})$ is open. □

We saw above that it is possible to find operators $S, T$ with $TS = I$ but neither invertible. Here $T$ is surjective though not injective, and it has a right-inverse. The existence of the right-inverse guarantees that $T$ is surjective; the converse is true for Hilbert spaces (Exercise 6.3.5) but not true in general. While we state this result here, its proof requires ideas from Section 6.3.

***6.2.4. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. The following statements are equivalent:*

*(i) $T$ admits a right-inverse;*

*(ii)* $\ker T$ *is complemented.*

*Proof.* Exercise 6.3.6. □

*6.2.5. Remark*. We can now produce a concrete example of a surjective map with no right inverse. Let $T : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})/c_0$ be the quotient map. Then $T$ is a bounded surjective map between Banach spaces, and $\ker T = c_0$. As $c_0$ is not complemented (Proposition 5.3.6), by Proposition 6.2.4 we have that $T$ admits no right inverse.

**6.2.1. Exercises.**

(6.2.1) Show that if $R$ is a ring with unit, and both $a$ and $ab$ are invertible, then $b$ is invertible. Using the algebra $\mathcal{B}(\ell^2(\mathbb{N}))$, show that it is possible to have $ab$ invertible with neither $a$ nor $b$ invertible.

(6.2.2) Let $S \in \mathcal{B}(\mathcal{X})$ with $\|I - S\| = 1$. Decide (and justify) whether such an $S$ is always invertible, sometimes invertible and sometimes not, or never invertible.

(6.2.3) Let $S \in \mathcal{B}(\mathcal{X})$ such that $\|I - S\| = \frac{1}{2}$. Decide (and justify) whether such an $S$ is always invertible, sometimes invertible and sometimes not, or never invertible.

(6.2.4) Let $S \in \mathcal{B}(\mathcal{X})$ such that $S^2 = S$. Decide (and justify) whether such an $S$ is always invertible, sometimes invertible and sometimes not, or never invertible.

## 6.3. Baire's Theorem and its Corollaries

We will make use, in an essential way, of the following very well-known result about metric spaces:

***6.3.1. Theorem. (The Baire Category Theorem)***

*Let $\{V_k\}$ be a countable family of **open dense** sets in a complete metric space $\mathcal{X}$. Then $\bigcap_k V_k$ is dense.*

*Proof.* Let $W \subset \mathcal{X}$ be a nonempty open subset. Since $V_1$ is dense, the open set $W \cap V_1$ is nonempty, so there exists $x_1 \in W \cap V_1$ and $\varepsilon_1$ with $0 < \varepsilon_1 < 1$ such that $\overline{B_{\varepsilon_1}(x_1)} \subset W \cap V_1$. Since $V_2$ is dense, there exists $x_2$ in the open set $B_{\varepsilon_1}(x_1) \cap V_2$ and so we can find $\varepsilon_2 < 1/2$ such that $\overline{B_{\varepsilon_2}(x_2)} \subset B_{\varepsilon_1}(x_1) \cap V_2$. Inductively, there exists $x_{k+1}$ in the open set $B_{\varepsilon_k}(x_k) \cap V_{k+1}$ and a closed ball $\overline{B_{\varepsilon_{k+1}}}(x_{k+1}) \subset B_{\varepsilon_k}(x_k) \cap V_{k+1}$, with $\varepsilon_{k+1} < 1/(k+1)$. This way we construct a sequence $\{x_k\}_k$. Now, given $\varepsilon > 0$, there exists $k_0 \in \mathbb{N}$ with $k_0 > 2/\varepsilon$. Then, if $k, h \geq k_0$, both $x_k$ and $x_h$ lie in the ball $B_{\varepsilon_{k_0}}(x_{k_0})$, so $d(x_k, x_h) < 2\varepsilon_{k_0} < 2/k_0 < \varepsilon$. This shows that the sequence $\{x_k\}$ is Cauchy, and by the completeness of $\mathcal{X}$ there exists a limit $x \in \mathcal{X}$; using that $\{x_{k+1}, x_{k+2}, \ldots\} \subset B_{\varepsilon_{k+1}}(x_{k+1})$ we have that $x$ is in the closed ball $\overline{B_{\varepsilon_{k+1}}(x_{k+1})}$; then $x \in \overline{B_{\varepsilon_{k+1}}(x_{k+1})} \subset B_{\varepsilon_k}(x_k) \cap V_{k+1}$ for all $k$. So $x \in V_k$ for all $k$, and $x \in W \cap V_1$ (going up in the chain of inclusions of balls). Then $W \cap (\bigcap_k V_k)$ is nonempty. As $W$ was an arbitrary open set, this shows that $\bigcap_k V_k$ is dense. □

*6.3.2. Remark* . The Category Theorem is a result about complete metric spaces (no vector space structure is used). The result is also valid for any locally compact Hausdorff topological space (Exercise 6.3.2; note that neither result can imply the other directly, for there are metric spaces which are not locally compact, and locally compact spaces which are not metrizable).

Closer to our current concerns, the Category Theorem works in any Banach space; and recall that if $\mathcal{X}$ is normed, then $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{C})$ and $\mathcal{B}(\mathcal{X})$ are Banach.

*6.3.3. Remark*. The Category Theorem is often stated and used, instead of in terms of countable intersections of open dense sets, in terms of countable unions of **nowhere dense** sets. These are sets (also called **rare**) such that the interior of their closure is empty; equivalently, the complement of their closure is open and dense (Exercise 6.3.1). A set is said to be **first category** or **meagre** if it is a countable union of nowhere dense sets, and **second category** if it is not of first category. The Category Theorem states that all complete metric spaces (and all locally compact Hausdorff spaces, as per Exercise 6.3.2) are second category sets.

Indeed, if $\mathcal{X} = \bigcup_k M_k$ where $M_k$ is nowhere dense for all $k$, then $\mathcal{X} = \bigcup_k \overline{M_k}$, and so $\bigcap_k (\mathcal{X} \setminus \overline{M_k}) = \varnothing$. Since $\mathcal{X} \setminus \overline{M_k}$ is open and dense for all $k$, the conclusion contradicts the Category Theorem. So a complete metric space (or a locally compact Hausdorff space) $\mathcal{X}$ is never a countable union of nowhere dense sets.

*6.3.4. Example.* A typical application of Baire's Category Theorem is to prove that a Hamel basis in a Banach space is necessarily uncountable (Exercise 6.3.3). Another typical application is the proof that if $\mathcal{X}$ is Banach, $T \in \mathcal{B}(\mathcal{X})$ and for each $x \in \mathcal{X}$ there exists $n \in \mathbb{N}$ with $T^n x = 0$, then there exists $m \in \mathbb{N}$ with $T^m = 0$ (Exercise 6.3.14).

Besides many minor applications like the ones mentioned above, the Baire Category Theorem is the basis for several crucial results that we will prove below. We begin with the Open Mapping Theorem. We have already seen examples of open maps, as linear functionals (Proposition 5.7.12) and quotient maps (Proposition 5.3.13) are open.

***6.3.5. Theorem. (The Open Mapping Theorem)***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be surjective. Then $T$ is an open map.*

*Proof.* Since any open set in $\mathcal{X}$ is a union of balls, it is enough to show that the image through $T$ of a ball is open (as the image of the union is the union of the images, Proposition 1.1.1). Because of the linearity of $T$ it is enough to show that $TB_1^{\mathcal{X}}(0)$ is open. Note that $\mathcal{X} = \bigcup_n B_n^{\mathcal{X}}(0)$. Since $T$ is surjective, its image is all of $\mathcal{Y}$, so

$$\mathcal{Y} = T(\mathcal{X}) = T\Big(\bigcup_{n\in\mathbb{N}} B_n^{\mathcal{X}}(0)\Big) = \bigcup_{n\in\mathbb{N}} TB_n^{\mathcal{X}}(0).$$

By the Category Theorem (as in Remark 6.3.3), there exists $n \in \mathbb{N}$ such that $\overline{TB_n^{\mathcal{X}}(0)}$ has nonempty interior. This means that there exist $y_0 \in \mathcal{Y}$ and $r > 0$ with $B_r^{\mathcal{Y}}(y_0) \subset \overline{TB_n^{\mathcal{X}}(0)}$.

Now let $y \in B_1^{\mathcal{Y}}(0)$. Then $y_0 + ry \in B_r^{\mathcal{Y}}(y_0)$; we deduce that

$$ry = (y_0 + ry) - y_0 \in \overline{TB_n^{\mathcal{X}}(0) - TB_n^{\mathcal{X}}(0)} = \overline{T(B_n^{\mathcal{X}}(0) - B_n^{\mathcal{X}}(0))} = \overline{TB_{2n}^{\mathcal{X}}(0)}.$$

Hence $y \in \overline{TB_\gamma^{\mathcal{X}}(0)}$, where $\gamma = 2n/r$. Thus $\overline{B_1^{\mathcal{Y}}(0)} \subset \overline{TB_\gamma^{\mathcal{X}}(0)}$. Given any $y \in \mathcal{Y}$ and $\varepsilon > 0$, since $y/\|y\| \in \overline{TB_\gamma^{\mathcal{X}}(0)}$ there exists $x' \in B_\gamma^{\mathcal{X}}(0)$ such that $\|y/\|y\| - Tx'\| < \varepsilon/\|y\|$. If we let $x = \|y\|\, x'$, we have

$$\|y - Tx\| < \varepsilon, \qquad \|x\| = \|y\|\,\|x'\| < \gamma\|y\|. \tag{6.2}$$

Now consider $y \in B_{\gamma^{-1}}^{\mathcal{Y}}(0) \subset \mathcal{Y}$. By (6.2), there exists $x_1$ with $\|x_1\| < \gamma\|y\| < 1$ and $\|y - Tx_1\| < \frac{1}{2\gamma}$. We continue inductively in the following way: given $x_k$ with

$$\|x_k\| < \frac{1}{2^{k-1}}, \qquad \|y - \sum_{j=1}^k Tx_j\| < \frac{1}{\gamma\, 2^k}, \tag{6.3}$$

we use (6.2) on $y - \sum_{j=1}^k Tx_j$ and $1/(\gamma\, 2^{k+1})$ to get $x_{k+1}$ with

$$\|x_{k+1}\| < \frac{1}{2^k}, \qquad \|y - \sum_{j=1}^k Tx_j - Tx_{k+1}\| < \frac{1}{\gamma\, 2^{k+1}},$$

so that (6.3) is now satisfied for $k+1$.

The construction implies that $y = \lim_k \sum_{j=1}^k Tx_j = \sum_{j=1}^\infty Tx_j$. Also

$$\Big\| \sum_{j=k+1}^\infty x_j \Big\| \leq \sum_{j=k+1}^\infty 2^{-j+1} = 2^{-k+1},$$

so the sequence of partial sums $\{\sum_{j=1}^k x_j\}$ is Cauchy and by completeness it converges to some $x \in \mathcal{X}$. The two convergences and the continuity of $T$ imply that $Tx = y$. Also,

$$\|x\| = \lim_{k\to\infty} \Big\| \sum_{j=1}^k x_j \Big\| \leq \sum_{j=1}^\infty \|x_j\| \leq \sum_{j=1}^\infty 2^{-j+1} = 2.$$

As $y \in B^{\mathcal{Y}}_{\gamma^{-1}}(0)$ was arbitrary, we have shown that $B^{\mathcal{Y}}_{\gamma^{-1}}(0) \subset TB^{\mathcal{X}}_2(0)$. Scaling we get that for any $r > 0$ there exists $r' > 0$ with $B^{\mathcal{Y}}_{r'}(0) \subset TB^{\mathcal{X}}_r(0)$.

Now, given any $y \in TB^{\mathcal{X}}_1(0)$, let $x \in B^{\mathcal{X}}_1(0)$ with $y = Tx$. Choose $\delta > 0$ with $\|x\| < \delta < 1$. Then $y \in TB^{\mathcal{X}}_\delta(0)$. By the above, there exists $r' > 0$ with $B^{\mathcal{Y}}_{r'}(0) \subset TB^{\mathcal{X}}_{1-\delta}(0)$. Then

$$B^{\mathcal{Y}}_{r'}(y) = y + B^{\mathcal{Y}}_{r'}(0) \subset TB^{\mathcal{X}}_\delta(0) + TB^{\mathcal{X}}_{1-\delta}(0) = TB^{\mathcal{X}}_1(0),$$

so $TB^{\mathcal{X}}_1(0)$ is open. □

Two important and frequently used consequences of the Open Mapping Theorem are the Inverse Mapping Theorem and the Closed Graph Theorem.

**6.3.6. Theorem. (The Inverse Mapping Theorem)**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ bijective. Then $T^{-1}$ is bounded.*

*Proof.* By the Open Mapping Theorem, for any $V \subset \mathcal{X}$ open, the pre-image through $T^{-1}$ is $TV$, which is open. So $T^{-1}$ is continuous. □

*6.3.7. Example.* Here is an example of the kind of things one can do with the Inverse Mapping Theorem. Let us use it to prove that in a finite-dimensional vector space, all norms are equivalent (we did this in Theorem 5.2.2). We will cheat a bit in the sense that to apply the Inverse Mapping Theorem we need to know that the spaces are complete, and the proof of completeness of finite-dimensional spaces came after showing that all norms are equivalent. But the point is simply to show what kind of information one gets from the Inverse Mapping Theorem. Because equivalence of norms is an equivalence relation, it is enough to show that an arbitrary norm is equivalent to a given norm. Let $\mathcal{X}$ be a finite dimensional normed vector space and let $\{e_1, \ldots, e_n\}$ be a basis for $x$. Put $c = \max\{\|e_n\| : \ n\}$ and define a norm $p$ on $\mathcal{X}$ by

$$p\Big(\sum_{j=1}^n a_j\, e_j\Big) = \sum_{j=1}^n |a_j|.$$

Let $T : \mathcal{X} \to \mathcal{X}$ be the identity operator, considered as map from $(\mathcal{X}, p)$ to $(\mathcal{X}, \|\cdot\|)$; it is obviously a bijection. We have, for $x = \sum_{j=1}^n a_j e_j$,

$$\|x\| = \Big\|T\Big(\sum_{j=1}^n a_j\, e_j\Big)\Big\| = \Big\|\sum_{j=1}^n a_j\, e_j\Big\| \le c\sum_{j=1}^n |a_j| = c\,p(x).$$

Then $T$ is a bounded bijection. By the Inverse Mapping Theorem (6.3.6), $T^{-1}$ is bounded. That is, as $T^{-1} : (\mathcal{X}, \|\cdot\|) \to (\mathcal{X}, p)$,

$$p(x) = p(T^{-1}x) \le \|T^{-1}\|\,\|x\|.$$

*6.3.8. Example*. Consider the space $c_{00}$ of complex sequences with finitely many nonzero entries, with the supremum norm, and let $T : c_{00} \to c_{00}$ be the operator

$$T : (a_1, a_2, \ldots) \longmapsto (a_1, a_2/2, a_3/3, \ldots). \tag{6.4}$$

Then $T$ is bounded (Exercise 6.1.11). As a function, $T$ is bijective; but its inverse

$$T^{-1} : (a_1, a_2, \ldots) \longmapsto (a_1, 2a_2, 3a_3, \ldots)$$

is not bounded (Exercise 6.3.7). Does this contradict the Inverse Mapping Theorem (6.3.6)? No, it doesn't, because $c_{00}$ is not Banach. If we complete $c_{00}$ we get $c_0$, the Banach space of complex sequences that converge to zero. The operator $T$, being bounded, can be extended to $c_0$ via a density argument (or just by noticing that the definition (6.4) works on $c_0$), but it is not bijective anymore (Exercise 6.1.12).

Here is something we promised long ago (in Remark 5.3.3, to be precise). The Inverse Mapping Theorem allows us to show that there is no need to distinguish between algebraic and topological direct sums in a Banach space.

**6.3.9. Proposition.**

*Let $\mathcal{X}$ be a Banach space, and $M, N \subset \mathcal{X}$ closed subspaces. If $\mathcal{X} = M \oplus N$ as an algebraic direct sum, then $\mathcal{X} = M \oplus_T N$ as a topological direct sum.*

*Proof.* Define $S : M \oplus N \to \mathcal{X}$ by $S(x, y) = x + y$. This map is a bijection by hypothesis. Also

$$\|S(x, y)\| = \|x + y\| \leq \|x\| + \|y\| = \|(x, y)\|,$$

so $S$ is bounded. By Theorem 6.3.6, its inverse $S^{-1}$ is also bounded, and so the direct sum is topological. □

*6.3.10. Remark.* We do need both spaces to be closed for Proposition 6.3.9 to work. For instance let $\mathcal{X} = \ell^\infty(\mathbb{N})$ and $N = c_0$. We know from Proposition 5.3.6 that $c_0$ is not complemented. But we can always take a Hamel basis of $N$ and extend it to a basis of $\mathcal{X}$; then the elements of the basis not in $N$ span a subspace $M$ and we have $\mathcal{X} = M + N$ as a direct sum. The fact that $N$ is not complemented guarantees—via contradicting Proposition 6.3.9—that $M$ is not closed.

**6.3.11. Definition.**

*Given $T : \mathcal{X} \to \mathcal{Y}$, the* ***graph of*** *$T$ is the set*

$$\mathcal{G}(T) = \{(x, Tx) : \ x \in \mathcal{X}\} \subset \mathcal{X} \oplus \mathcal{Y},$$

*that we consider as a normed vector space with the norm $\|(x, y)\| = \|x\| + \|y\|$*

The above is not the only choice for a norm in the direct sum, since we could have $\left(\|x\|^2 + \|y\|^2\right)^{1/2}$, or $\max\{\|x\|, \|y\|\}$, or others; but all these reasonable norms are equivalent, so they generate the same topology.

We have that $\mathcal{G}(T)$ is closed if and only if $x_k \to 0$ and $Tx_k \to y$ together imply that $y = 0$ (Exercise 6.3.9).

**6.3.12. Theorem.** ***(The Closed Graph Theorem)***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and let $T : \mathcal{X} \to \mathcal{Y}$ be linear. Then $\mathcal{G}(T)$ is closed if and only if $T$ is bounded.*

*Proof.* If $T$ is bounded and $\{(x_j, Tx_j)\}_j$ is Cauchy in $\mathcal{G}(T)$, then it converges to $(x, y) \in \mathcal{X} \oplus \mathcal{Y}$. This in particular implies that $x_j \to x$, $Tx_j \to y$. Since $T$ is continuous, $Tx = y$, and $\mathcal{G}(T)$ is closed.

Conversely, suppose that $\mathcal{G}(T)$ is closed. Let $\pi_1 : \mathcal{G}(T) \to \mathcal{X}$ be the projection $\pi_1(x, Tx) = x$. Then $\pi_1$ is linear and bounded. It is clearly surjective, and because its domain is a graph, it is injective. So by the Inverse Mapping Theorem 6.3.6, $\pi_1^{-1}$ is bounded. The map $\pi_2 : \mathcal{G}(T) \to \mathcal{Y}$ given by $\pi_2(x, Tx) = Tx$ is also linear and bounded. Now $T = \pi_2 \circ \pi_1^{-1}$ is bounded. $\square$

We have proven that the Open Mapping Theorem implies the Inverse Mapping Theorem, and that the Inverse Mapping Theorem implies the Closed Graph Theorem. It is also possible to prove that the Closed Graph Theorem implies the Open Mapping Theorem (Exercise 6.3.8), so the three results are logically equivalent.

*6.3.13. Example*. The hypothesis that $\mathcal{X}$ is Banach is essential in Theorem 6.3.12. Recall for instance the example from page 370: let $\mathcal{X} = \mathcal{Y} = \mathbb{C}[x]$ with the norm $\|p\|_\infty = \sup\{|p(t)| : \ t \in [0,1]\}$. Let $T : \mathcal{X} \to \mathcal{Y}$ be given by $Tp = p'$. Then $T$ is unbounded, as was shown, and its graph is closed. Indeed, if $p_j \to 0$ uniformly and $p_j' \to g$ uniformly, then (using the uniform convergence)

$$\int_0^x g = \lim_j \int_0^x p_j' = \lim_j p_j(x) - p_j(0) = 0.$$

As this can be done for any $x$, we get that $g = 0$, so $\mathcal{G}(T)$ is closed. The fact that the derivative is often unbounded as a linear operator makes a compelling case for one to consider unbounded operators. And indeed, unbounded operators appear frequently in physics and in the study of differential equations, among other areas. The unbounded operators that are "good" (that is, that they are more or less tractable) are precisely those with closed graph. That said, there are natural examples of unbounded operators without closed graph, and that cannot be extended to an operator with closed graph. Let $\mathcal{X} = L^2[0,1]$ with its dense subspace $\mathcal{X}_0 = C[0,1]$. Define $T : \mathcal{X}_0 \to \mathcal{X}_0$ by $Tf = f(0)$. This is unbounded, with the same proof as in Exercise 5.6.17. If we take $f_n = (1-nt)\,1_{[0,\frac{1}{n}]}$ then $f_n \in C[0,1]$ for all $n$, $\|f_n\|_2 \to 0$, and $Tf_n = 1$ for all $n$, so $\mathcal{G}(T)$ is not closed (and it cannot be made closed by extending $T$ to a larger subspace).

We can use the Closed Graph Theorem to obtain a variation of Proposition 6.3.9.

***6.3.14. Proposition.***

*Let $\mathcal{X}$ be a Banach space, and $M, N \subset \mathcal{X}$ closed subspaces with $\mathcal{X} = M \oplus N$. Then the projection $P_M$ onto $M$ along $N$ is bounded.*

*Proof.* Suppose that $x_j \to x$ and $P_M x_j \to y$. Writing $x_j = z_j + w_j$, with $z_j \in M$ and $w_j \in N$, we have $P_M x_j = z_j$. So $z_j \to y$. As $M$ is closed, $y \in M$. Since $x = y + (x - y)$ and $x - y = \lim(z_j + w_j) - z_j = \lim w_j$, we get that $w_j \to x - y$ and so $x - y \in N$. Thus $P_M x = y$. This shows that the graph of $P_M$ is closed, and so $P_M$ is bounded by the Closed Graph Theorem. □

As another example of how the Closed Graph Theorem can be used, consider this:

***6.3.15. Proposition.***

*Let $\mathcal{X}$ be a Banach space, and $T : \mathcal{X} \to \mathcal{X}^*$ be linear and such that $(Tx)x \geq 0$ for all $x \in \mathcal{X}$. Then $T$ is bounded.*

A quick comment about this result. The expression $(Tx)x$ looks fairly unusual, but it is actually very common in the appropriate context. Namely, when $\mathcal{X}$ is a Hilbert space, the above reads $\langle Tx, x\rangle \geq 0$, which is to say that $T$ is positive (or, as said in linear algebra, positive-semidefinite). In other words, the proposition shows that a positive-semidefinite operator is necessarily bounded.

*Proof.* Suppose that $x_n \to x$ and $Tx_n \to \phi$. If we show that $\phi = Tx$, then the Closed Graph Theorem guarantees that $T$ is bounded.

Given $z \in \mathcal{X}$,

$$0 \leq (T(x_n - z))(x_n - z) = (Tx_n)x_n + (Tz)z - (Tx_n)z - (Tz)x_n.$$

Taking limit over $n$, and noting that convergent sequences in a norm space are bounded (for the double limit in the first term),

$$0 \leq \phi(x) + (Tz)z - \phi(z) - (Tz)x = \phi(x - z) - (Tz)(x - z).$$

So $(Tz)(x - z) \leq \phi(x - z)$. Note that $z$ was arbitrary; if $w = x - z$, then $(T(x - w)w \leq \phi(w)$, which we write as

$$(Tx - \phi)w \leq (Tw)w, \qquad w \in \mathcal{X}.$$

In particular, $(Tx - \phi)w \in \mathbb{R}$ since $(Tw)w \in \mathbb{R}$. Given any $w \in \mathcal{X}$, the inequality also works for $\pm w/n$, which gives us

$$\pm n(Tx - \phi)w \leq (Tw)w, \qquad w \in \mathcal{X},\ n \in \mathbb{N}.$$

This forces $(Tx - \phi)w = 0$. As this occurs for all $w \in \mathcal{X}$, we have shown that $Tx = \phi$. □

Another huge consequence of the Baire Category Theorem is the **Uniform Boundedness Principle**:

**6.3.16. Theorem. (Banach–Steinhaus; Uniform Boundedness Principle)**

*Let $\mathcal{X}$ be a Banach space and $\mathcal{Y}$ a normed space. Let $\mathcal{F} \subset \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be a collection of operators such that $\sup\{\|Tx\| : T \in \mathcal{F}\} < \infty$ for each $x \in \mathcal{X}$. Then*

$$\sup\{\|T\| : T \in \mathcal{F}\} < \infty.$$

*Proof.* Let $\mathcal{X}_n = \{x \in \mathcal{X} : \|Tx\| \leq n \text{ for all } T \in \mathcal{F}\}$. Then $\mathcal{X} = \bigcup_n \mathcal{X}_n$. Since $\mathcal{X}$ is complete, by the Baire Category Theorem there exists $n$ such that the closure of $\mathcal{X}_n$ has nonempty interior. But $\mathcal{X}_n$ is already closed, so there exists $\delta > 0$ and $x \in \mathcal{X}_n$ such that $\overline{B_\delta(x)} \subset \mathcal{X}_n$. Now let $z \in \mathcal{X}$ with $\|z\| \leq \delta$. Then $z + x \in \overline{B_\delta(x)}$ and, for any $T \in \mathcal{F}$,

$$\|Tz\| \leq \|T(z+x)\| + \|-Tx\| = \|T(z+x)\| + \|Tx\| \leq n + n = 2n.$$

If $z \in \mathcal{X}$ with $\|z\| = 1$, then

$$\|Tz\| = \frac{1}{\delta}\,\big\|T(\delta 2\, z)\big\| \leq \frac{2n}{\delta}.$$

This shows that $\|T\| \leq 4n/\delta$, for all $T \in \mathcal{F}$. □

The corollary below should not be confused with Proposition 6.1.5, where the key condition for the completeness of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is the completeness of $\mathcal{Y}$ and, more importantly, the convergence is uniform. Note that for the argument to work below it is essential to be dealing with sequences and not nets.

**6.3.17. Corollary.**

*Let $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ is Banach and $\mathcal{Y}$ is normed. Suppose that $\{T_n x\}$ converges for each $x \in \mathcal{X}$. Then $Tx = \lim T_n x$ defines a bounded operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

*Proof.* The fact that $T$ is a linear function is immediate. For each $x \in \mathcal{X}$, since $\{T_n x\}$ is a convergent sequence in a normed space, it is bounded. That is, for each $x$

$$\sup\{\|T_n x\| : n \in \mathbb{N}\} < \infty.$$

By the Uniform Boundedness Principle (Theorem 6.3.16), there exists $k > 0$ with $\|T_n\| \leq k$ for all $n$. Then, for any $x \in \mathcal{X}$,

$$\|Tx\| \leq \|Tx - T_n x\| + \|T_n x\| \leq \|Tx - T_n x\| + k\|x\|.$$

As this holds for all $n$ and $Tx - T_n x \to 0$, we get that $\|Tx\| \leq k\|x\|$. □

We will see a weaker version of Corollary 6.3.17 in Corollary 7.1.26. It is important to mention that Corollary 6.3.17 does not show nor it implies that $T_n \to T$ in norm. For example let $\mathcal{X} = \ell^2(\mathbb{N})$ with $\{e_n\}$ the canonical basis, and let $T_n$ be the orthogonal projection onto $\operatorname{span}\{e_1, \ldots, e_n\}$. Then $T_n x \to x$ for all $x \in \mathcal{X}$, but $\|I - T_n\| = 1$ for all $n$.

*6.3.18. Remark*. There is an elementary proof of the Uniform Boundedness Principle (Theorem 6.3.16) that does not use Baire's Category Theorem, due to Alan Sokal [Sok11]. This should not be a great surprise, as the proof of Baire's Category Theorem is in itself also rather elementary.

**6.3.1. Exercises.**

(6.3.1) Let $\mathcal{X}$ be a metric space. Show that $A \subset \mathcal{X}$ is nowhere dense if and only if $\mathcal{X} \setminus \overline{A}$ is open and dense.

(6.3.2) Prove that Baire's Category Theorem 6.3.1 holds for locally compact Hausdorff topological spaces (*Hint: use the finite intersection property instead of completeness*).

(6.3.3) Show that any Hamel basis of an infinite-dimensional Banach space is uncountable (*Hint: show that finite-dimensional subspaces are nowhere dense*).

(6.3.4) Show an example of an infinite-dimensional normed space that is a countable union of nowhere dense subsets.

(6.3.5) Let $\mathcal{H}$ be a Hilbert space. Prove that $T \in \mathcal{B}(\mathcal{H})$ is surjective if and only if it admits a right inverse $S \in \mathcal{B}(\mathcal{H})$. The same assertion is not true for Banach spaces (Remark 6.2.5).

(6.3.6) Let $\mathcal{X}$ be a Banach space. Show that $T \in \mathcal{B}(\mathcal{X})$ admits a right inverse if and only if $T$ is surjective and $\ker T$ is complemented.

(6.3.7) Show that $T^{-1}$ in Example 6.3.8 is unbounded.

(6.3.8) Prove that the Closed Graph Theorem 6.3.12 implies the Open Mapping Theorem 6.3.5.

(6.3.9) Complete the proof of Theorem 6.3.12 by showing that for a linear map $T : \mathcal{X} \to \mathcal{Y}$ with $\mathcal{X}, \mathcal{Y}$ Banach, the graph $\mathcal{G}(T)$ is closed if and only if for every sequence $\{x_k\} \subset \mathcal{X}$ and $y \in \mathcal{Y}$, if $x_k \to 0$ and $Tx_k \to y$ then $y = 0$.

(6.3.10) Let $\mathcal{X}$ be a Banach space and $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ closed subspaces such that $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_1 \subset \mathcal{X}_1$ and $T\mathcal{X}_2 \subset \mathcal{X}_2$. Show that $T$ is invertible if and only if $T|_{\mathcal{X}_1} \in \mathcal{B}(\mathcal{X}_1)$ and $T|_{\mathcal{X}_2} \subset \mathcal{B}(\mathcal{X}_2)$ are invertible.

(6.3.11) Let $\mathcal{X}$ be a Banach space and $T : \mathcal{X} \to \mathcal{X}^*$ be a linear map. Use the Closed Graph Theorem to show that if $(Tx)y = (Ty)x$ for all $x, y \in \mathcal{X}$, then $T$ is bounded.

(6.3.12) Let $\mathcal{X}$ be a Banach space and $T : \mathcal{X} \to \mathcal{X}^*$ be a linear map. Use the Uniform Boundedness Principle to show that if $(Tx)y = (Ty)x$ for all $x, y \in \mathcal{X}$, then $T$ is bounded.

(6.3.13) *(Compare with Exercise 6.3.14)* Show an example of a normed space $\mathcal{X}$ and $T \in \mathcal{B}(\mathcal{X})$ such that for every $x \in \mathcal{X}$ there exists $n \in \mathbb{N}$ with $T^n x = 0$, but such that $T^m \neq 0$ for all $m \in \mathbb{N}$.

(6.3.14) *(Compare with Exercise 6.3.13)* Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$. Assume that, for each $x \in \mathcal{X}$, there exists $n \in \mathbb{N}$ such that $T^n x = 0$. Show that there exists $m \in \mathbb{N}$ such that $T^m = 0$.

CHAPTER

# 7

# Weak topologies

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two topologies on a set $\mathcal{X}$. Recall from subsection 1.8.2 that if $\mathcal{T}_1 \subset \mathcal{T}_2$, we say that $\mathcal{T}_1$ is **weaker** (or **coarser**) than $\mathcal{T}_2$, and that $\mathcal{T}_2$ is **stronger** (or **finer**) than $\mathcal{T}_1$. We will say by convention that any topology is weaker (and stronger) than itself.

Since the set $\mathcal{P}(\mathcal{X})$ of all subsets of $\mathcal{X}$ is a topology, and since an intersection of topologies is a topology, for any given family of subsets $\mathcal{F} \subset \mathcal{P}(\mathcal{X})$ there is a smallest topology that contains $\mathcal{F}$, denoted here by $\operatorname{Top} \mathcal{F}$. We think of $\operatorname{Top} \mathcal{F}$ as the **topology generated by** $\mathcal{F}$.

## 7.1. The weak topology

Let $\mathcal{X}$ be a locally convex space. Then we can consider its dual $\mathcal{X}^*$. By definition, every functional in $\mathcal{X}^*$ is continuous. Now, is the topology of $\mathcal{X}$ the weakest that makes all functionals in $\mathcal{X}^*$ continuous? The answer is rarely yes, as we will see.

One central reason to care about weaker topologies, is that the weaker the topology (that is, the less open sets) the easier it is to get compactness—since there are less open covers to care about. And compactness is good, it allows one to prove nice things, as we will see over and over again. Continuity is harder for a weaker topology, though. In any case, the point is that there are a number of topologies that we can define on a topological vector space, and each is useful for certain things. In this section and the next one we will discuss the most common alternative topologies, namely the weak and weak* topologies.

**7.1.1. Definition.**

*The **weak topology** on a topological vector space $\mathcal{X}$ is the topology*

$$\sigma(\mathcal{X},\mathcal{X}^*) = \text{Top}\{\varphi^{-1}(V) : \ \varphi \in \mathcal{X}^*, \ V \subset \mathbb{C}, \ \textit{open}\ \}$$

*For $X \subset \mathcal{X}$, we will denote its closure in the $\sigma(\mathcal{X},\mathcal{X}^*)$-topology by $\overline{X}^{\sigma(\mathcal{X},\mathcal{X}^*)}$.*

*When a net $\{x_j\}$ converges to $x$ weakly, we write $x_j \xrightarrow{\text{weak}} x$.*

The definition implies that the weak topology is the weakest topology such that every $\varphi \in \mathcal{X}^*$ is continuous (Exercise 7.1.2). By definition, since every $\varphi \in \mathcal{X}^*$ is continuous in the original topology, $\sigma(\mathcal{X},\mathcal{X}^*)$ is weaker than the natural topology of $\mathcal{X}$.

If $\mathcal{X}$ is bad enough then $\sigma(\mathcal{X},\mathcal{X}^*)$ can be very uninteresting; when $\mathcal{X}^* = \{0\}$—like in Proposition 5.6.17—we have that $\sigma(\mathcal{X},\mathcal{X}^*) = \{\varnothing,\mathcal{X}\}$. But the spaces we pay attention to are locally convex, and so their duals are big and the weak topology is non-trivial.

***7.1.2. Lemma.***

*If $\mathcal{X}$ is any TVS such that $\mathcal{X}^*$ separates points (in particular, if $\mathcal{X}$ is a LCS), then the topology $\sigma(\mathcal{X},\mathcal{X}^*)$ is Hausdorff.*

*Proof.* Let $x, y \in \mathcal{X}$ with $x \neq y$. Since $\mathcal{X}^*$ separates points (when $\mathcal{X}$ is locally convex, by Theorem 5.7.18 applied to $\{x\}$ and $\{y\}$), there exists $\varphi \in \mathcal{X}^*$ with $\varphi(x) \neq \varphi(y)$. Let $\varepsilon = |\varphi(x) - \varphi(y)|/3$, and put

$$V_x = \varphi^{-1}(B_\varepsilon(\varphi(x)), \quad V_y = \varphi^{-1}(B_\varepsilon(\varphi(y)).$$

Then $x \in V_x$, $y \in V_y$, and $V_x \cap V_y = \varnothing$ because $B_\varepsilon(\varphi(x)) \cap B_\varepsilon(\varphi(y)) = \varnothing$. □

Here is a first step in getting a good handle of $\sigma(\mathcal{X},\mathcal{X}^*)$. Note the change from preimages of arbitrary open sets to preimages of balls. This change entails no technical difficulty (Exercise 7.1.3).

***7.1.3. Lemma.***

*Let $\mathcal{X}$ be a TVS. Define*

$$\mathcal{I} = \left\{ \bigcap_{j=1}^m \varphi_j^{-1}(B_{r_j}(\lambda_j)) : m \in \mathbb{N}, \ \lambda_j \in \mathbb{C}, \ r_j > 0, \ \varphi_j \in \mathcal{X}^*, \ j = 1,\ldots,m \right\}.$$

*Then*

$$\sigma(\mathcal{X},\mathcal{X}^*)=\Big\{\bigcup_{W\in\mathcal{W}}W:\ \mathcal{W}\subset\mathcal{I}\Big\}.$$

*Proof.* The set $\tilde{\mathcal{I}}=\{\bigcup_{W\in\mathcal{W}}W:\ \mathcal{W}\subset\mathcal{I}\}$ is contained in any topology that contains $\sigma(\mathcal{X},\mathcal{X}^*)$, so it is contained in $\sigma(\mathcal{X},\mathcal{X}^*)$. It also contains the sets $\varphi^{-1}(V)$ that form a generating set for $\sigma(\mathcal{X},\mathcal{X}^*)$ (since any open subset of $\mathbb{C}$ is a union of balls). So it is enough to show that $\tilde{\mathcal{I}}$ is a topology. The empty set is $\varphi^{-1}(\varnothing)$ for any $\varphi\in\mathcal{X}^*$, so it is in $\mathcal{I}$. If we have a family $\{\bigcup_{W\in\mathcal{W}_j}W\}_j$, then

$$\bigcup_j\bigcup_{W\in\mathcal{W}_j}W=\bigcup_{W\in\mathcal{W}}W,\qquad\text{where }\mathcal{W}=\bigcup_j\mathcal{W}_j.$$

Similarly, if we have two sets $\{\bigcup_{W\in\mathcal{W}_j}W\}_{j=1}^2\}$ then, using that $\mathcal{I}$ is by definition closed under finite intersections,

$$\Big(\bigcup_{W\in\mathcal{W}_1}W\Big)\cap\Big(\bigcup_{W\in\mathcal{W}_2}W\Big)=\bigcup_{W_1\in\mathcal{W}_1}\bigcup_{W_2\in\mathcal{W}_2}W_1\cap W_2\in\Big\{\bigcup_{W\in\mathcal{W}}W:\ \mathcal{W}\subset\mathcal{I}\Big\}.$$

and we easily get intersections of finite families by induction. □

While there will be a number of occasions where the use of neighbourhoods will be unavoidable, in many cases we will deal with a more concrete description of weak convergence:

**7.1.4. Proposition.**

*Let $\{x_j\}$ be a net in $\mathcal{X}$. Then the following statements are equivalent:*

*(i)* $x_j\xrightarrow{\text{weak}}x$;

*(ii) for every* $\varphi\in\mathcal{X}^*$, $\lim_j\varphi(x_j)=\varphi(x)$.

*Proof.* Assume that $x_j\xrightarrow{\text{weak}}x$. Let $\varphi\in\mathcal{X}^*$, $\varepsilon>0$; then the preimage $\varphi^{-1}(B_\varepsilon(\varphi(x)))$ is a weak-open neighbourhood of $x$. Since $x_j\xrightarrow{\text{weak}}x$, there exists $j_0$ such that, for any $j>j_0$, $x_j\in\varphi^{-1}(B_\varepsilon(\varphi(x)))$, i.e. $|\varphi(x_j)-\varphi(x)|<\varepsilon$. So $\varphi(x_j)\to\varphi(x)$.

Conversely, suppose that $\varphi(x_j)\to\varphi(x)$ for all $\varphi\in\mathcal{X}^*$, and assume that $W$ is an open weak-neighbourhood of $x$ in $\mathcal{X}$. By Lemma 7.1.3 there exist $\varphi_1,\ldots,\varphi_m\in\mathcal{X}^*$ and balls $B_{\varepsilon_k}(c_k)$, with $\varepsilon_1,\ldots,\varepsilon_m>0$ and $c_1,\ldots,c_m\in\mathbb{C}$, such that

$$x\in\bigcap_{k=1}^m\varphi_k^{-1}(B_{\varepsilon_k}(c_k))\subset W.$$

This means that $|\varphi_k(x) - c_k| < \varepsilon_k$ for all $k$. As $\varphi_k(x_j) \to \varphi_k(x)$ for all $k = 1, \dots, m$, we can find $j_0$ such that, for all $j \geq j_0$, $|\varphi_k(x_j) - \varphi_k(x)| < \min\{\varepsilon_k - |\varphi_k(x) - c_k| : k = 1, \dots, m\}$. This guarantees that, when $j \geq j_0$,

$$|\varphi_k(x_j) - c_k| \leq |\varphi_k(x_j) - \varphi_k(x)| + |\varphi_k(x) - c_k| < \varepsilon_k, \qquad k = 1, \dots m.$$

Thus $x_j \in \bigcap_{k=1}^m \varphi_k^{-1}(B_{\varepsilon_k}(c_k)) \subset W$ for all $j \geq j_0$ and we have shown that $x_j \xrightarrow{\text{weak}} x$. □

We record here an interesting consequence of Proposition 7.1.4:

***7.1.5. Proposition.***

*If the TVS $\mathcal{X}$ is considered as a TVS with its weak topology, its dual is still $\mathcal{X}^*$.*

*Proof.* We denote $\mathcal{X}^w = (\mathcal{X}, \sigma(\mathcal{X}, \mathcal{X}^*))^*$. Since the weak topology $\sigma(\mathcal{X}, \mathcal{X}^*)$ is defined in such a way that every $\varphi \in \mathcal{X}^*$ is $\sigma(\mathcal{X}, \mathcal{X}^*)$-continuous, we have $\mathcal{X}^* \subset \mathcal{X}^w$. Now suppose that $\varphi : \mathcal{X} \to \mathbb{C}$ is linear and $\sigma(\mathcal{X}, \mathcal{X}^*)$-continuous; if $x_j \to x$, then $x_j \xrightarrow{\text{weak}} x$ (since the weak topology is weaker than the original topology), so $\varphi(x_j) \to \varphi(x)$; thus $\varphi$ is continuous and so $\varphi \in \mathcal{X}^*$, showing that $\mathcal{X}^w \subset \mathcal{X}^*$. □

*7.1.6. Example.* We can use Proposition 7.1.4 to show a familiar explicit example where the weak topology is strictly weaker than the norm topology—this is always the case for infinite-dimensional spaces. Let $\mathcal{H} = \ell^2(\mathbb{N})$. Write $\{\xi_j\}$ for the canonical basis for $\mathcal{H}$. Then $\|\xi_j\| = 1$, and $\|\xi_j - \xi_k\| = \sqrt{2}$ if $j \neq k$, making the sequence $\{\xi_j\}$ not convergent (it doesn't even have convergent subsequences).

On the other hand, by the Riesz Representation Theorem 4.5.4 or by Proposition 5.6.3 we know that every functional in $\mathcal{H}^*$ is given by inner product with a vector from $\mathcal{H}$. Given $\xi \in \mathcal{H}$, we know that

$$\|\xi\|^2 = \sum_j |\langle \xi, \xi_j \rangle|^2$$

(Parseval's equality (4.13)). So $\langle \xi_j, \xi \rangle \to 0$ for each $\xi \in \mathcal{H}$, i.e. $\xi_j \xrightarrow{\text{weak}} 0$.

*7.1.7. Remark.* Proposition 7.1.4 simplifies our dealings with the weak topology in a couple ways:

(i) it remarks the fact that the weak topology is linear, i.e. $x_j \xrightarrow{\text{weak}} x$ if and only if $x_j - x \xrightarrow{\text{weak}} 0$;

(ii) it implies that the weak topology is locally convex for any TVS: indeed, it is given by the family of seminorms $\{p_\varphi : \varphi \in \mathcal{X}^*\}$, where $p_\varphi(x) = |\varphi(x)|$ (Exercise 7.1.8).

While arbitrary open sets in $\sigma(\mathcal{X}, \mathcal{X}^*)$ are not easy to handle, it is not difficult to see, using Lemma 7.1.3, that $\sigma(\mathcal{X}, \mathcal{X}^*)$ admits a local base at 0 of neighbourhoods

$$N(\varphi_1, \dots, \varphi_k; \varepsilon) = \{x \in \mathcal{X} : |\varphi_j(x)| < \varepsilon, \ j = 1, \dots, k\}, \tag{7.1}$$

indexed by $k \in \mathbb{N}$, $\varphi_1, \dots, \varphi_k \in \mathcal{X}^*$, $\varepsilon > 0$ (Exercise 7.1.4). Moreover, we may restrict to those tuples $\varphi_1, \dots, \varphi_k$ that are linearly independent (Exercise 7.1.5). We also note that since $\mathcal{X}^*$ is a vector space, the families $\{N(\varphi_1, \dots, \varphi_n; \varepsilon) : \varphi_1, \dots, \varphi_n \in \mathcal{X}^*, \ \varepsilon > 0\}$ and $\{N(\varphi_1, \dots, \varphi_n; 1) : \varphi_1, \dots, \varphi_n \in \mathcal{X}^*\}$ are equal.

We showed in Example 7.1.6 that the weak topology is strictly weaker than the norm topology on a Hilbert space. As it turns out, this happens in any infinite-dimensional normed space:

***7.1.8. Proposition.***

*Let $\mathcal{X}$ be an infinite-dimensional normed space. Then the weak topology is strictly weaker than the original topology.*

*Proof.* We will show that the unit open ball is not open in the weak topology. Consider any weak open neighbourhood $V$ of $0 \in \mathcal{X}$. Then there exists a basic neighbourhood $N(\varphi_1, \dots, \varphi_k; 1) \subset V$. As $\mathcal{X}$ is infinite-dimensional, $M = \ker \varphi_1 \cap \dots \cap \ker \varphi_k \neq \{0\}$ (Proposition 5.5.12). As $M \subset N(\varphi_1, \dots, \varphi_k; 1)$ is a subspace, $N(\varphi_1, \dots, \varphi_k; 1)$ contains elements of arbitrarily large norm. It follows that that the unit ball does not contain any neighbourhood of 0, and so it is not open in the weak topology. ☐

*7.1.9. Remark* . We cannot expect Proposition 7.1.8 to extend to arbitrary locally convex spaces. Indeed, we showed in Proposition 7.1.5 that the dual of $(\mathcal{X}, \sigma(\mathcal{X}, \mathcal{X}^*))$ is again $\mathcal{X}^*$, and so we have scores of examples of locally convex spaces where the weak topology agrees with the original topology.

*7.1.10. Remark.* Let us stress the fact, from the proof of Proposition 7.1.8, that in an infinite-dimensional normed space any weak-neighbourhood of 0 contains a non-zero subspace.

In the next result we will use the natural identification of $\mathcal{X}$ as a subspace of $\mathcal{X}^{**}$ (the dual of the dual), by having $\hat{x}(\varphi) = \varphi(x)$ for $x \in \mathcal{X}$, $\varphi \in \mathcal{X}^*$. This identification is isometric, because as an element of $\mathcal{X}^{**}$,

$$\begin{aligned} \|\hat{x}\|_{\mathcal{X}^{**}} &= \sup\{|\hat{x}(\varphi)| : \ \varphi \in \mathcal{X}^*, \ \|\varphi\| = 1\} \\ &= \sup\{|\varphi(x)| : \ \varphi \in \mathcal{X}^*, \ \|\varphi\| = 1\} = \|x\|_{\mathcal{X}}, \end{aligned} \tag{7.2}$$

where the last equality is due to the Hahn–Banach Theorem (Corollary 5.7.7). It is common to use the notation $\varphi(x) = \langle x, \varphi\rangle$—inspired by the Hilbert space case, see Example 7.1.6—which presents the duality $\varphi(x) = \hat{x}(\varphi)$ in a very clear way.

**7.1.11. *Proposition.***

*A weakly convergent sequence on a normed space $\mathcal{X}$ is bounded.*

*Proof.* Let $\{x_n\}$ be a sequence with $x_n \xrightarrow{\text{weak}} x$. Recall that $\mathcal{X}^*$ is Banach (Proposition 5.5.8). For each $\varphi \in \mathcal{X}^*$, $\sup\{|\hat{x}_n(\varphi)| : \ n \in \mathbb{N}\} < \infty$, since any convergent sequence in $\mathbb{C}$ is bounded. But then, by the Uniform Boundedness Principle (Theorem 6.3.16), applied to $\{\hat{x}_n\} \subset \mathcal{B}(\mathcal{X}^*, \mathbb{C}) = \mathcal{X}^{**}$, we get $\sup\{\|x_n\| : n \in \mathbb{N}\} = \sup\{\|\hat{x}_n\| : n \in \mathbb{N}\} < \infty$. $\square$

Before discussing several things about Proposition 7.1.11, let us note that with the same technique we can prove the following:

**7.1.12. *Proposition.***

*A weakly compact set on a normed space $\mathcal{X}$ is bounded.*

*Proof.* Exercise 7.1.9. $\square$

*7.1.13. Remark*. The word "sequence" is essential in Proposition 7.1.11, and it is not only about countability. That's because convergent nets need not be bounded, **even if the index set is countable**: consider, for instance, the net $\{x_n\}_{n\in\mathbb{Z}} \subset \mathbb{R}$ given by

$$x_n = \begin{cases} -n & \text{if } n < 0 \\ \frac{1}{n} & \text{if } n > 0 \end{cases}.$$

Then $x_n \to 0$, but $\sup\{|x_n| : n \in \mathbb{Z}\} = \infty$.

If the above example feels like cheating (because the unbounded part is not "involved" in the convergence), here is a more meaty example. Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space with orthonormal basis $\{\xi_n\}$ and consider the set $R = \{\sqrt{n}\,\xi_n : \ n \in \mathbb{N}\}$. Then

- $0 \in \overline{R}^{\sigma(\mathcal{H},\mathcal{H}^*)}$, so there exists a net $\{\sqrt{n_j}\,\xi_{n_j}\}$ (indexed by weak open neighbourhoods of 0) with $\sqrt{n_j}\,\xi_{n_j} \to 0$ weakly; by Proposition 7.1.11 this net cannot be a sequence, since it is unbounded. To see that $0 \in \overline{R}^{\sigma(\mathcal{H},\mathcal{H}^*)}$, note that if $0 \notin \overline{\{\sqrt{n}\,\xi_n\}}^{\sigma(\mathcal{H},\mathcal{H}^*)}$, then there exists an open neighbourhood
$$W = \{\xi : \ |\varphi_j(\xi)| < 1,\ j = 1, \ldots, N\}$$
such that $W \cap \overline{\{\sqrt{n}\,\xi_n\}}^{\sigma(\mathcal{H},\mathcal{H}^*)} = \varnothing$. Via the Riesz Representation Theorem (Theorem 4.5.4) this means that there exist $\eta_1, \ldots, \eta_N \in \mathcal{H}$ and indices $j_n \in \{1, \ldots, N\}$ with
$$1 \leq |\langle \sqrt{n}\,\xi_n, \eta_{j_n}\rangle|, \qquad n \in \mathbb{N}$$
(because at least one $|\langle \sqrt{n}\,\xi_n, \eta_j\rangle| < 1$ has to fail for $\sqrt{n}\xi_n$ not to be in $W$). Then
$$\sum_{j=1}^{N} \|\eta_j\|^2 = \sum_{j=1}^{N}\sum_{n=1}^{\infty} |\langle \xi_n, \eta_j\rangle|^2 = \sum_{n=1}^{\infty}\sum_{j=1}^{N} |\langle \xi_n, \eta_j\rangle|^2 \geq \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$
a contradiction. So $0 \in \overline{\{\sqrt{n}\,e_n\}}^{\sigma(\mathcal{H},\mathcal{H}^*)}$.
- No subsequence of $R$ converges weakly to 0. Indeed, if $\{\eta_m\}_{m\in\mathbb{N}} \subset R$, then either there are infinitely many distinct elements, making the sequence unbounded and thus not convergent by Proposition 7.1.11; or there are only finitely many distinct elements, which allows us to choose a constant subsequence, which again does not converge to 0.
- The weak topology of $\mathcal{H}$ is not metrizable. If the weak topology in $\mathcal{H}$ were metrizable, there would be a countable basis of neighbourhoods around 0. This would allow us to choose a subsequence of $R$ that converges weakly to 0, which we saw is not possible.
- A similar argument can be used to construct an unbounded weakly-convergent sequence in $\ell^p(\mathbb{N})$, for $p \in (1, \infty)$ (Exercise 7.1.20).

The last point in Remark 7.1.13 is not a particular feature of Hilbert spaces, but rather a general phenomenon on normed spaces:

***7.1.14. Proposition.***

*Let $\mathcal{X}$ be an infinite-dimensional normed space. Then the weak topology on $\mathcal{X}$ is not metrizable.*

*Proof.* We will show that if the weak topology is metrizable, then $\dim \mathcal{X} < \infty$. If the weak topology is metrizable, there exists a countable basis for the weak topology at 0. This implies that there exists a countable family of basic neighbourhoods that generate the basis at 0. These are of the form

$$N_k = \{x \in \mathcal{X} : \ |\varphi_{k,j}(x)| < 1, \ j = 1, \ldots, m(k)\}.$$

Let $\varphi \in \mathcal{X}^*$. Since $V = \{x \in \mathcal{X} : \ |\varphi(x)| < 1\}$ is an open set containing 0, there exists $k$ such that $N_k \subset V$. If $\bigcap_{j=1}^{m(k)} \ker \varphi_{k,j} = \{0\}$, then $\dim \mathcal{X} < \infty$ by Proposition 5.5.12 and there is nothing to prove. Otherwise let $x \in \bigcap_{j=1}^{m(k)} \ker \varphi_{k,j}$ with $x \neq 0$. As $N_k \subset V$, we get that $|\varphi(x)| < 1$. For any $n \in \mathbb{N}$, we also have $nx \in \bigcap_{j=1}^{m(k)} \ker \varphi_{k,j}$; then $|\varphi(nx)| < 1$, which is to say $|\varphi(x)| < \frac{1}{n}$. As this can be done for any $n \in \mathbb{N}$, we have shown that $\varphi(x) = 0$. In other words,

$$\bigcap_{j=1}^{m(k)} \ker \varphi_{k,j} \subset \ker \varphi.$$

By Lemma 5.5.10, $\varphi \in \operatorname{span}\{\varphi_{k,1}, \ldots, \varphi_{k,m(k)}\}$. Thus the countable set

$$\bigcup_{k=1}^{\infty} \{\varphi_{k,1}, \ldots, \varphi_{k,m(k)}\}$$

spans $\mathcal{X}^*$. As $\mathcal{X}^*$ is complete (Proposition 5.5.8), if infinite-dimensional then any basis would be uncountable (Exercise 6.3.3); therefore $\dim \mathcal{X}^* < \infty$. By Lemma 5.7.8, $\dim \mathcal{X} < \infty$. □

On the other hand, we have

**7.1.15. Proposition.**

*Let $\mathcal{X}$ be a normed space. If $\mathcal{X}^*$ is separable, then $B_1^{\mathcal{X}}(0)$ is weak-metrizable.*

*Proof.* Exercise 7.1.15. □

The fact that the weak topology is generally not metrizable allows one to consider the **weak sequential closure** of a subset of a normed space, namely the set of limits of sequences in the space. See Exercise 7.1.17 for a concrete example of how the weak sequential closure can differ from the weak closure.

While we saw above that the weak topology is not generally metrizable, compactness can be decided on sequences. This is the Eberlein–Šmulian Theorem (Theorem 7.2.28); while the result is about the weak topology, its proof requires the weak*-topology, hence its placement in the next section.

Next is a nice implication from convexity:

***7.1.16. Theorem.***

*If $K \subset \mathcal{X}$ is a convex subset of the locally convex space $\mathcal{X}$, then*
$$\overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)} = \overline{K}.$$

*Proof.* Since $\sigma(\mathcal{X},\mathcal{X}^*)$ is weaker than the original topology, $\overline{K} \subset \overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)}$. Now fix $x_0 \in \overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)} \setminus \overline{K}$. We apply Theorem 5.7.18 (second form of geometric Hahn–Banach) to the compact set $\{x_0\}$ and the closed convex set $\overline{K}$ to obtain $\varphi \in \mathcal{X}^*$ and $c \in \mathbb{R}$ with
$$\operatorname{Re}\varphi(x_0) < c < \operatorname{Re}\varphi(x), \ x \in \overline{K}.$$
Then $\overline{K} \subset C = \{x \in \mathcal{X} : \ \operatorname{Re}\varphi(x) \geq c\}$. This last set $C$ is weakly closed (Exercise 7.1.7), so $\overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)} \subset C$. In particular, $x_0 \in C$, a contradiction. Thus $\overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)} = \overline{K}$. □

***7.1.17. Corollary. (Mazur's Lemma)***

*Let $\mathcal{X}$ be a locally convex space, and $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ with $x_n \xrightarrow{\text{weak}} x$. Then there exists a net $\{x'_m\}_m$, with $x'_m \to x$ and such that each $x'_m$ is a convex combination of some $x_n$. If $\mathcal{X}$ is metrizable, then we can choose the $\{x'_m\}$ to be a sequence.*

*Proof.* Exercise 7.1.10. □

**7.1.1. Some examples.** We follow up a bit on the examples from Section 5.6.

Recall that $c_{00}$ is the set of all sequences with finite support and that $c_{00}$ is dense in $\ell^p(\mathbb{N})$ (Exercise 2.8.9).

***7.1.18. Lemma.***

*Let $p \in (1,\infty]$, and $\{f_n\}_{n\in\mathbb{N}} \subset \ell^p(\mathbb{N})$ be a bounded sequence. Then the following statements are equivalent:*

*(i) $\langle f_n, g\rangle \to 0$ for all $g \in \ell^q(\mathbb{N})$;*

*(ii) $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$.*

*Proof.* The implication (i) $\Longrightarrow$ (ii) is immediate, since $c_{00} \subset \ell^q(\mathbb{N})$.

Conversely, let $g \in \ell^q(\mathbb{N})$. Since the sequence $\{f_n\}$ is bounded, there exists $M > 0$ with $\|f_n\|_p < M$ for all $n$. Let $\varepsilon > 0$. As $q < \infty$, there exists

$g' \in c_{00}$ with $\|g' - g\|_q < \varepsilon/M$. We have, using Hölder's inequality,

$$0 \leq |\langle f_n, g\rangle| \leq |\langle f_n, g'\rangle| + |\langle f_n, g' - g\rangle| \leq |\langle f_n, g'\rangle| + \varepsilon.$$

Hence $0 \leq \limsup_n |\langle f_n, g\rangle| \leq \varepsilon$. By the Limsup Routine, the limit exists and is 0. □

*7.1.19. Remark*. Neither the boundedness nor the $p > 1$ conditions can be dropped in Lemma 7.1.18 (Exercises 7.1.13 and 7.1.14).

From Lemma 7.1.18 we obtain

***7.1.20. Proposition.***

*Let $p \in [1, \infty]$. A weakly convergent sequence in $\ell^p(\mathbb{N})$ is pointwise convergent. If $1 < p \leq \infty$, a bounded pointwise convergent sequence is weakly convergent.*

*Proof.* If $\{f_n\}$ is weakly convergent to 0, we have in particular

$$0 = \lim_n \langle f_n, e_k\rangle = \lim_n f_n(k).$$

Conversely, if $1 < p \leq \infty$ and if the bounded sequence $\{f_n\}$ converges pointwise to 0, then $\langle f_n, g\rangle \to 0$ for any $g \in c_{00}$. By Lemma 7.1.18, $\{f_n\}$ is weakly convergent. □

*7.1.21. Remark*. The result in Proposition 7.1.20 does not hold in arbitrary $L^p$ spaces. For example the bounded sequence $\{g_n\} \subset L^p[0, 1]$ given by

$$g_n(t) = \operatorname{sgn}(\sin(\pi n t))$$

converges weakly to 0, but $g_n \neq 0$ a.e. (Exercise 7.1.19).

In the case of $\ell^1(\mathbb{N})$ we have the following, which together with Proposition 7.1.8 stresses the fact that one needs nets and not just sequences when working with the weak topology.

***7.1.22. Proposition. (Schur's Property)***

*A sequence in $\ell^1(\mathbb{N})$ converges weakly if and only if it converges in norm.*

*Proof.* Norm convergence always implies weak convergence, so there is only one implication to care about.

Let $\{x_n\} \subset \ell^1(\mathbb{N})$ be a sequence. We write $x_n = (x_{n,k})$. Since both topologies are linear we may assume without loss of generality that $x_n \xrightarrow{\text{weak}} 0$. This means that

$$\sum_k x_{n,k} y_k \xrightarrow[n\to\infty]{} 0, \qquad y \in \ell^\infty(\mathbb{N}).$$

By Proposition 7.1.20, $\lim_n x_{n,k} = 0$ for all $k$.

Suppose that $\{x_n\}$ does not converge in norm to 0. By passing to a subsequence if necessary, we may assume that there exists $\delta > 0$ with $\|x_n\|_1 \geq \delta$ for all $n$. This allows us to normalize the sequence to have $\|x_n\|_1 = 1$ for all $n$, while still $x_n \xrightarrow{\text{weak}} 0$ (Exercise 7.1.12). Next choose $n_1$ and $m_1$ such that

$$\sum_{j=n_1}^{\infty} |x_{1,j}| < \frac{1}{5} \qquad \sum_{j=1}^{n_1} |x_{m_1,j}| < \frac{1}{5}.$$

The first inequality is possible due to $x_1 \in \ell^1(\mathbb{N})$, and the second one—after $n_1$ has been determined—because of the pointwise convergence to 0. Inductively, we choose $n_k$ and $m_k$ with $n_k > n_{k-1} + 1$, $m_k > m_{k-1} + 1$, and

$$\sum_{j=n_k}^{\infty} |x_{m_{k-1},j}| < \frac{1}{5} \qquad \sum_{j=1}^{n_k} |x_{m_k,j}| < \frac{1}{5}.$$

We have, for all $k$ (recall that $\|x_n\|_1 = 1$),

$$\sum_{j=n_k+1}^{n_{k+1}} |x_{m_k,j}| = 1 - \sum_{j=1}^{n_k} |x_{m_k,j}| - \sum_{j=n_{k+1}+1}^{\infty} |x_{m_k,j}| > 1 - \frac{1}{5} - \frac{1}{5} = \frac{3}{5}.$$

Now define $y \in \ell^\infty(\mathbb{N})$ by letting $\gamma(0) = 0$, $\gamma(z) = \bar{z}/|z|$ (so that $z\gamma(z) = |z|$) and

$$y_j = \gamma(x_{m_k,j}), \qquad \text{if } n_k < j \leq n_{k+1}.$$

Then $\|y\|_\infty = 1$, so $y \in \ell^\infty(\mathbb{N})$, and

$$\begin{aligned}
|\langle x_{m_k}, y\rangle| &= \Big| \sum_{j=1}^{n_k} x_{m_k,j} y_j + \sum_{j=n_k+1}^{n_{k+1}} x_{m_k,j} y_j + \sum_{j=n_{k+1}+1}^{\infty} x_{m_k,j} y_j \Big| \\
&\geq \Big| \sum_{j=n_k+1}^{n_{k+1}} x_{m_k,j} y_j \Big| - \sum_{j=1}^{n_k} |x_{m_k,j}| - \sum_{j=n_{k+1}+1}^{\infty} |x_{m_k,j}| \\
&= \sum_{j=n_k+1}^{n_{k+1}} |x_{m_k,j}| - \sum_{j=1}^{n_k} |x_{m_k,j}| - \sum_{j=n_{k+1}+1}^{\infty} |x_{m_k,j}| \\
&\geq \frac{3}{5} - \frac{2}{5} = \frac{1}{5}.
\end{aligned}$$

This implies that $\{x_{m_k}\}$ does not converge weakly to zero, a contradiction. So a weakly convergent sequence in $\ell^1(\mathbb{N})$ converges in norm. $\square$

*7.1.23. Remark.* It is worth stressing the fact that Proposition 7.1.22 does not say that weak convergence and norm convergence are the same in $\ell^1(\mathbb{N})$. The result is about *sequences*, where in general we need nets. Typically a convergent net need not be bounded, and that would differentiate it from a convergent sequence. But in this case, there are even *bounded* nets in $\ell^1(\mathbb{N})$ that converge weakly but not in norm. Here is an example. Let $\mathcal{J}$ be the set of finite partitions of $\mathbb{N}$ that contain no singletons. That is, the elements of $\mathcal{J}$ are of the form $(A_1,\ldots,A_n)$, where $n\in\mathbb{N}$, $A_j\subset\mathbb{N}$ for all $j$, $|A_j|\geq 2$, pairwise disjoint, and $\mathbb{N}=\bigcup_{j=1}^n A_j$. For uniqueness we require that $A_1,\ldots,A_n$ are ordered in such a way that if $j<k$ then $\min A_j<\min A_k$. We order $\mathcal{J}$ by saying that $(A_1,\ldots,A_n)\leq(B_1,\ldots,B_m)$ if $m\geq n$ and for each $j\in\{1,\ldots,n\}$ there exists $k(j)$ such that $B_j\subset A_{k(j)}$; that is, a partition $B$ is bigger than an partition $A$ if $B$ is a refinement of $A$. For each $A\in\mathcal{J}$ we define $x_A\in\ell^1(\mathbb{N})$ by writing $A_1=\{a_1,a_2,\ldots,a_r\}$ in the natural order, and putting

$$x_A=\frac{1}{2}\,1_{\{a_1\}}-\frac{1}{2}\,1_{\{a_2\}}.$$

We have $\|x_A\|_1=\frac{1}{2}+\frac{1}{2}=1$ for all $A\in\mathcal{J}$. Now fix $y\in\ell^\infty(\mathbb{N})$ and $\varepsilon>0$. By definition of $\ell^\infty(\mathbb{N})$, the set $\{y_n\}_n\subset\mathbb{C}$ is bounded; then its closure is compact and so we can cover it with finitely many open balls of radius $\varepsilon$, say $B_\varepsilon(z_1),\ldots,B_\varepsilon(z_s)$. Let $A_j'=\{k\in\mathbb{N}:\ y_k\in B_\varepsilon(z_j)\}$, $j=1,\ldots,s$, and put $A_1=A_1'$, $A_j=A_j'\setminus\bigcup_{h<j}A_h$. Then, after reordering if needed, $A=(A_1,\ldots,A_s)\in\mathcal{J}$. Write $A_1=\{a_1,\ldots,a_r\}$. For any $B\in\mathcal{J}$ with $A\leq B$, we have that $b_1,b_2$ lie in the same $A_j$, and so

$$|y_{b_1}-y_{b_2}|\leq|y_{b_1}-z_j|+|z_j-y_{b_2}|<\varepsilon+\varepsilon=2\varepsilon.$$

Then

$$|\langle x_B,y\rangle|=\frac{1}{2}\,|y_{b_1}-y_{b_2}|<2\varepsilon.$$

This shows that $x_B\to 0$ weakly, while $\|x_B\|_1=1$ for all $B$. So the net $\{x_B\}_{B\in\mathcal{J}}$ is bounded, converges weakly to zero, and it does not converge in norm.

There is no contradiction with Proposition 7.1.15 because the dual of $\ell^1(\mathbb{N})$, which is $\ell^\infty(\mathbb{N})$, is not separable.

**7.1.24. Definition.**

*Let $\mathcal{X}$ be a topological vector space. A set $S\subset\mathcal{X}$ is said to be* ***weakly bounded*** *if*

$$\sup\{|\varphi(x)|:\ x\in S\}<\infty$$

*for each $\varphi\in\mathcal{X}^*$.*

For an arbitrary topological vector space this definition may not amount to much, as we may have $\mathcal{X}^* = \{0\}$.

**7.1.25. Proposition.**

*A weakly bounded set in a normed space is bounded.*

*Proof.* A careful examination of the proof of Proposition 7.1.11 will show that the fact that $\{x_n\}$ is a sequence was only used to determine that it was weakly bounded. Hence the proof carries on with the weakly bounded set $S$ instead of the sequence. $\square$

We can now weaken the hypothesis in Corollary 6.3.17:

**7.1.26. Corollary.**

*Let $\{T_n\} \subset \mathcal{B}(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ is Banach and $\mathcal{Y}$ is normed. Suppose that the sequence $\{T_n x\}$ converges weakly for each $x \in \mathcal{X}$. Then $Tx = \lim T_n x$ defines a bounded operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

*Proof.* The fact that $T$ is a linear function is immediate. For each $x \in \mathcal{X}$ and $\varphi \in \mathcal{Y}^*$, since $\{\varphi(T_n x)\}$ is a convergent sequence in $\mathbb{C}$, it is bounded. That is, the set $\{T_n x : \ n\} \subset \mathcal{Y}$ is weakly bounded. By Proposition 7.1.25,

$$\sup\{\|T_n x\| : \ n \in \mathbb{N}\} < \infty.$$

Now we can repeat the rest of the proof of Corollary 6.3.17. $\square$

With similar ideas we can prove the following.

**7.1.27. Proposition.**

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, and $T : \mathcal{X} \to \mathcal{Y}$ be linear. If $T$ is continuous when both $\mathcal{X}$ and $\mathcal{Y}$ are considered with their weak topologies, then $T$ is bounded.*

*Proof.* Suppose that $T$ is unbounded. Then there exists a sequence $\{x_n\} \subset \mathcal{X}$ with

$$\|x_n\| = 1, \qquad \|Tx_n\| \geq n^2, \qquad n \in \mathbb{N}.$$

Let $x_n' = x_n/n$. Then $x_n' \to 0$. By hypothesis, $Tx_n' \to 0$ weakly; that is, for each $\psi \in \mathcal{Y}^*$ we have $\psi(Tx_n') \to 0$. Hence the sequence $\{Tx_n'\}$ is weakly bounded; so bounded by Proposition 7.1.25, say $\|Tx_n'\| \leq c$ for all $n$. This gives us $n \leq \|Tx_n'\| \leq c$ for all $n$, a contradiction. So $T$ is bounded. $\square$

**7.1.2. Exercises.**

(7.1.1) Prove that if $\mathcal{T}_1 \subset \mathcal{T}_2$ are topologies on a set $\mathcal{X}$, with $\mathcal{T}_1$ Hausdorff and $\mathcal{T}_2$ compact, then $\mathcal{T}_1 = \mathcal{T}_2$ (*Hint: consider the identity map* $(\mathcal{X}, \mathcal{T}_2) \to (\mathcal{X}, \mathcal{T}_1)$).

(7.1.2) Let $\mathcal{X}$ be a TVS. Show that the weak topology $\sigma(\mathcal{X}, \mathcal{X}^*)$ is the weakest topology such that every $\varphi \in \mathcal{X}^*$ is continuous.

(7.1.3) Prove that

$$\sigma(\mathcal{X}, \mathcal{X}^*) = \mathrm{Top}\{\varphi^{-1}(B_r(\lambda)) : \ \varphi \in \mathcal{X}^*, \ r > 0, \ \lambda \in \mathbb{C}\}.$$

(7.1.4) Show that the sets $N(\varphi_1, \ldots, \varphi_k; \varepsilon)$, indexed by a positive integer $k \in \mathbb{N}$ and $\varphi_1, \ldots, \varphi_k \in \mathcal{X}^*$, $\varepsilon > 0$, form a local base for $\sigma(\mathcal{X}, \mathcal{X}^*)$ at 0.

(7.1.5) Show that the sets $N(\varphi_1, \ldots, \varphi_k; \varepsilon)$, where $\varphi_1, \ldots, \varphi_k \in \mathcal{X}^*$ are linearly independent and $\varepsilon > 0$, form a local base for $\sigma(\mathcal{X}, \mathcal{X}^*)$ at 0.

(7.1.6) Let $\mathcal{X}$ be a normed space and $\{x_n\}$ a weakly convergent net with $x_n \xrightarrow{\text{weak}} x$. Show that $\|x\| \le \liminf_n \|x_n\|$. Find an example where the inequality is strict.

(7.1.7) Use Proposition 7.1.4 to show that if $\varphi \in \mathcal{X}^*$, $c \in \mathbb{R}$, then $\{x \in \mathcal{X} : \operatorname{Re}\varphi(x) \ge c\}$ is weakly closed.

(7.1.8) Prove that the weak topology is a locally convex topology given by the seminorms $p_\varphi(x) = |\varphi(x)|$, $\varphi \in \mathcal{X}^*$.

(7.1.9) Prove Proposition 7.1.12.

(7.1.10) Prove Corollary 7.1.17.

(7.1.11) Show that the unit ball of $c_0$ is not weakly compact.

(7.1.12) Let $\mathcal{X}$ be a normed space and $\{x_n\} \subset \mathcal{X}$ such that $\|x_n\| \ge \delta > 0$ for all $n$, and such that $x_n \xrightarrow{\text{weak}} 0$. Show that $x_n/\|x_n\| \xrightarrow{\text{weak}} 0$.

(7.1.13) Show that the boundedness requirement in Lemma 7.1.18 cannot be dispensed with. That is find an example, for $p \in (1, \infty]$, of a sequence $\{f_n\} \subset \ell^p(\mathbb{N})$ such that $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$ and $h \in \ell^q(\mathbb{N})$ such that $\langle f_n, h\rangle$ does not converge to 0.

(7.1.14) Show that the $p > 1$ requirement in Lemma 7.1.18 cannot be dispensed with. That is, find an example of a bounded sequence $\{f_n\} \subset \ell^1(\mathbb{N})$ such that $\langle f_n, g\rangle \to 0$ for all $g \in c_{00}$ and $h \in \ell^\infty(\mathbb{N})$ such that $\langle f_n, h\rangle$ does not converge to 0.

(7.1.15) Let $\mathcal{X}$ be a normed space, with $\mathcal{X}^*$ separable. Show that the weak topology is metrizable on the unit ball $B_1^{\mathcal{X}}(0)$. *(Hint: if in need of inspiration, look at the proof of Corollary 7.2.20)*

(7.1.16) Let $\mathcal{X}$ be a Banach space and $K \subset \mathcal{X}$ convex. Show that $K$ is weakly closed if and only if $K \cap \overline{B_r(0)}$ is weakly closed for all $r > 0$.

(7.1.17) Let $\mathcal{H}$ be a separable Hilbert space and fix an orthonormal basis $\{\xi_n\}$. Let $S \subset \mathcal{H}$ be

$$S = \{\xi_m + m\xi_n : \ n, m \in \mathbb{N}\}.$$

Let $\overline{S}^{\text{w-seq}}$ denote the weak sequential closure of $S$. Prove that the inclusion $\overline{S}^{\text{w-seq}} \subset \overline{\overline{S}^{\text{w-seq}}}^{\text{w-seq}}$ is proper by showing that

$$0 \in \overline{\overline{S}^{\text{w-seq}}}^{\text{w-seq}} \setminus \overline{S}^{\text{w-seq}}.$$

(7.1.18) With $S$ as in Exercise 7.1.17, show that $0 \in \overline{S}^{\sigma(\mathcal{H},\mathcal{H}^*)}$.

(7.1.19) Let $p \in (1, \infty]$. Consider the sequence $\{g_n\} \subset L^p[0,1]$ given by

$$g_n(t) = \operatorname{sgn}(\sin(\pi n t)).$$

Show that $\{g_n\}$ converges weakly to 0, but $\{g_n\}$ does not converge pointwise to 0.

(7.1.20) Generalize the idea in Remark 7.1.13 to construct an example of a weakly-convergent unbounded net in $\ell^p(\mathbb{N})$ for $p \in [1, \infty)$.

## 7.2. The weak* topology

Since the dual $\mathcal{X}^*$ of a normed space $\mathcal{X}$ is itself a normed space, one can consider its dual, the **double dual** $\mathcal{X}^{**}$, which we have already mentioned. In particular, there is an isometric embedding (i.e. an isometric homomorphism, see (7.2)) of $\mathcal{X}$ into $\mathcal{X}^{**}$ via the duality $\hat{x}(\varphi) = \varphi(x)$ (or, in the alternative notation, $\langle \varphi, x\rangle = \langle x, \varphi\rangle$). Now, besides its own weak topology $\sigma(\mathcal{X}^*, \mathcal{X}^{**})$, we can also consider on $\mathcal{X}^*$ the weakest topology that makes the functionals coming from $\mathcal{X}$ continuous. As in general the inclusion $\mathcal{X} \subset \mathcal{X}^{**}$ is proper, this gives yet another topology, denoted by $\sigma(\mathcal{X}^*, \mathcal{X})$—a bit of reflection should make it clear that the notation makes sense—and called the **weak*-topology**.

More explicitly,

**7.2.1. Definition.**

*The **weak*-topology** (pronounced "weak-star-topology") on a topological vector space $\mathcal{X}$ is the topology*

$$\sigma(\mathcal{X}^*, \mathcal{X}) = \operatorname{Top}\{\hat{x}^{-1}(W) : \ x \in \mathcal{X},\ W \subset \mathbb{C} \ \text{ open}\}$$

*For $Y \subset \mathcal{X}^*$, we will denote its closure in the $\sigma(\mathcal{X}^*, \mathcal{X})$-topology by $\overline{Y}^{\sigma(\mathcal{X}^*,\mathcal{X})}$.*

*When a net $\{\varphi_j\}$ converges to $\varphi$ weakly, we will use the notation $\varphi_j \xrightarrow{\text{weak}^*} \varphi$.*

We remark that $\sigma(\mathcal{X}^*, \mathcal{X})$ makes sense for any topological vector space; all the definition needs is for $\mathcal{X}^*$ to exist. The definition of the weak*-topology makes sense even on the algebraic dual of an arbitrary real/complex vector space, though it is hard to see what use that could have. In practice, the dual can be fairly pathological for an arbitrary TVS (see, for instance, Propositions 5.6.17 and 5.6.21), so we will always consider the case $\mathcal{X}$ where is locally convex. When $\mathcal{X}$ is normed and the double dual exists, $\sigma(\mathcal{X}^*, \mathcal{X})$ is by definition weaker than $\sigma(\mathcal{X}^*, \mathcal{X}^{**})$. It is still Hausdorff for all locally convex spaces, however:

**7.2.2. Lemma.**

*Let $\mathcal{X}$ be a locally convex space. Then the topology $\sigma(\mathcal{X}^*, \mathcal{X})$ is Hausdorff.*

*Proof.* Exercise 7.2.3. ☐

Although the more abstract point of view is valuable and needed for many results, weak*-convergence is nothing but pointwise convergence:

**7.2.3. Proposition.**

*Let $\mathcal{X}$ be a topological space and $\{\varphi_j\}$ be a net in $\mathcal{X}^*$. Then the following statements are equivalent:*

*(i)* $\varphi_j \xrightarrow{\text{weak}^*} \varphi$;

*(ii) for every* $x \in \mathcal{X}$, $\lim_j \varphi_j(x) = \varphi(x)$.

*Proof.* Exercise 7.2.4. ☐

*7.2.4. Remark.* In analogy with the weak topology (see Remark 7.1.7), Proposition 7.2.3 implies that

(i) the weak* topology is linear, i.e. $\varphi_j \xrightarrow{\text{weak}^*} \varphi$ if and only if $\varphi_j - \varphi \xrightarrow{\text{weak}^*} 0$;

(ii) a local base at 0 for the weak* topology is given by the sets
$$N(x_1, \ldots, x_k; \varepsilon) = \{\varphi \in \mathcal{X}^* : \ |\varphi(x_j)| < \varepsilon, \ j = 1, \ldots, k\},$$
indexed by $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathcal{X}$, $\varepsilon > 0$.

(iii) If $\dim \mathcal{X} = \infty$, every weak*-neighbourhood of 0 is unbounded. Indeed, given $N(x_1, \ldots, x_k; \varepsilon)$, by Proposition 5.5.12 we have $0 \neq \bigcap_j \ker \hat{x}_j \subset N(x_1, \ldots, x_k; \varepsilon)$; as this is a subspace, it is unbounded. In particular, the weak*-topology cannot agree with the norm topology unless $\dim \mathcal{X} < \infty$.

(iv) The $\sigma(\mathcal{X}^*, \mathcal{X})$ topology is locally convex, given by the seminorms $\{p_x : x \in \mathcal{X}\}$, where
$$p_x(\varphi) = |\varphi(x)|.$$

A consequence of Propositions 7.2.3 and 7.1.4 that will be useful is this:

***7.2.5. Corollary.***

*If $\mathcal{X}$ is a normed space, then the isometric embedding $\iota : \mathcal{X} \to \mathcal{X}^{**}$ is a homeomorphism from $(\mathcal{X}, \sigma(\mathcal{X}, \mathcal{X}^*)) \to (\iota(\mathcal{X}), \sigma(\mathcal{X}^{**}, \mathcal{X}^*))$.*

*Proof.* From Proposition 7.2.3 and Proposition 7.1.4,
$$x_j \xrightarrow{\text{weak}} 0 \iff \varphi(x_j) \to 0 \text{ for all } \varphi \in \mathcal{X}^* \iff \hat{x}_j(\varphi) \to 0 \text{ for all } \varphi \in \mathcal{X}^*$$
$$\iff \hat{x}_j \xrightarrow{\text{weak}^*} 0.$$ □

Continuing with the analogy with the weak topology, we can prove the following. Recall that the dual of a normed space is a Banach space (Proposition 5.5.8).

***7.2.6. Proposition.***

*Let $\mathcal{X}$ be a Banach space. Then any weak* convergent sequence in $\mathcal{X}^*$ is bounded.*

*Proof.* Exercise 7.2.5. □

*7.2.7. Examples.* Let $\mathcal{X} = c_0$. We saw in Exercise 5.6.2 and Proposition 5.6.3 that $\mathcal{X}^* = \ell^1(\mathbb{N})$, $\mathcal{X}^{**} = \ell^\infty(\mathbb{N})$, where in both cases the duality is given by $\langle x, y\rangle = \sum_j x_j y_j$ (note that always one sequence will be summable and the other will be bounded).

Let $e_1, e_2, \ldots$ denote the canonical basis $e_j(k) = \delta_{j,k}$. These elements belong to all three spaces, and we have the following facts:

(i) In $c_0$, $e_j \xrightarrow{\text{weak}} 0$, while $\|e_j\| = 1$ for all $j$. So we have an example of a weakly convergent sequence which is not convergent (i.e. the weak topology is strictly weaker than the norm topology).

(ii) In $\ell^1(\mathbb{N})$, $e_j \xrightarrow{\text{weak}^*} 0$, but $e_j$ does not converge weakly (i.e. the weak* topology is strictly weaker than the weak topology).

(iii) Let $K = \operatorname{conv}\{e_j : j\} \subset \ell^1(\mathbb{N})$. Then 0 belongs to the weak* closure of $K$ but not to the weak closure of $K$. So the closures of a convex set in the weak and weak* topologies do not necessarily agree (compare with Theorem 7.1.16).

*7.2.8. Remark*. We mention here something that in some cases can lead to confusion, and it also shows that the requirement that $\mathcal{X}$ is Banach in Proposition 7.2.6 is essential. If $\mathcal{X}$ is a normed space that is not complete, then $\mathcal{X}^*$ is the dual both of $\mathcal{X}$ and of its completion. This opens the possibility of considering two different weak*-topologies on $\mathcal{X}^*$, namely the ones induced by $\mathcal{X}$ and by its completion. And these two topologies are usually different! Here is an example. Let $\mathcal{X} = c_{00}$, but considered as a (dense) subspace of $\ell^1(\mathbb{N})$, so we use the one-norm on $\mathcal{X}$. For both $\mathcal{X}$ and its completion $\ell^1(\mathbb{N})$ the dual is $\ell^\infty(\mathbb{N})$. Let $w_k \in \ell^\infty(\mathbb{N})$ be given by $w_k = \sum_{j \geq k} k e_j$. Then $w_k \to 0$ weak*, since any sequence in $\mathcal{X}$ is eventually zero, but

$$\sum_n \frac{w_k(n)}{n^2} = \sum_{n \geq k} \frac{k}{n^2} \geq k \int_k^\infty \frac{1}{x^2}\, dx = 1,$$

showing that $w_k$ does not converge to zero in the weak*-topology induced by the completion of $\mathcal{X}$. The distinction made by this example will be emphasized by Proposition 7.2.10

***7.2.9. Proposition.***

*Let $\mathcal{X}$ be an infinite-dimensional Banach space. Then $(\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$ is not metrizable.*

*Proof.* Suppose for contradiction that the topology is metrizable. Then the topology admits a countable local base of balls at 0. We may assume that the balls satisfy $B_1 \supset B_2 \supset \cdots \ni 0$. Since each ball contains a basic neighbourhood of 0, we deduce that there exists a local base $\{V_n\}$ at 0, where $V_n = \{\psi \in \mathcal{X}^* : |\psi(x_j)| < 1,\ j = 1, \ldots, n\}$ for a sequence $\{x_n\} \subset \mathcal{X}$. Given $x \in \mathcal{X}$, consider the neighbourhood $V_x = \{\psi \in \mathcal{X}^* : |\psi(x)| < 1\}$ of 0. As $\{V_n\}$ is a base, there exists $n$ such that $V_n \subset V_x$. Let

$$\psi \in \bigcap_{j=1}^n \ker \hat{x}_j \subset V_n \subset V_x.$$

As this is a subspace, we get that $m\psi \in V_x$ for all $m \in \mathbb{N}$. That is, $|\psi(x)| < \frac{1}{m}$ for all $m$, so $\psi(x) = 0$. That is,

$$\bigcap_{j=1}^n \ker \hat{x}_j \subset \ker \hat{x}.$$

By Lemma 5.5.10, we have that $\hat{x} \in \operatorname{span}\{\hat{x}_1, \ldots, \hat{x}_n\}$ and then, since $\mathcal{X}^*$ separates points, $x \in \operatorname{span}\{x_1, \ldots, x_n\}$. This shows that the countable set $\{x_k\}$ is a Hamel basis for $\mathcal{X}$, contradicting Exercise 6.3.3. $\square$

Notwithstanding Proposition 7.2.9, the closed unit ball of $(\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$ is metrizable when $\mathcal{X}$ is separable (Corollary 7.2.20).

Now we depart from the analogy with the weak topology. In particular, as seen in Examples 7.2.7, there is no analog to Theorem 7.1.16 for the weak*-topology.

**7.2.10. Proposition.**

*Let $\mathcal{X}$ be a locally convex space. Then the weak*-continuous functionals on $\mathcal{X}^*$ are precisely the point evaluations $\varphi \longmapsto \varphi(x)$. In consequence, the dual of $(\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$ is $\mathcal{X}$.*

*Proof.* Any point evaluation $\hat{x} : \varphi \longmapsto \varphi(x)$ is a weak*-continuous linear functional: if $\varphi_j \xrightarrow{\text{weak}^*} \varphi$, then by definition $\varphi_j(x) \to \varphi(x)$, which we see as $\hat{x}(\varphi_j) \to \hat{x}(\varphi)$; so $\hat{x}$ is a continuous linear functional; that is, $\hat{x} \in (\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))^*$.

We now check that every weak*-continuous linear functional is of the form $\hat{x}$. Let us fix $\psi \in (\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))^*$. By Proposition 5.5.9, and since the seminorms for $\sigma(\mathcal{X}^*, \mathcal{X})$ in $\mathcal{X}^*$ are given by the points in $\mathcal{X}$, there exist $c_1, \ldots, c_n > 0$ and $x_1, \ldots, x_n \in \mathcal{X}$ with

$$|\psi(\varphi)| \leq \sum_{j=1}^n c_j \, |\hat{x}_j(\varphi)|, \qquad \varphi \in \mathcal{X}^*.$$

This implies that $\bigcap_{j=1}^n \ker \hat{x}_j \subset \ker \psi$. Then Lemma 5.5.10 shows that there exist $a_1, \ldots, a_n \in \mathbb{C}$ with

$$\psi = \sum_{j=1}^n a_j \hat{x}_j = \widehat{\sum_{j=1}^n a_j x_j}.$$

It is straightforward to check that the map $x \longmapsto \hat{x}$ is linear, which justifies the last equality above. $\square$

*7.2.11. Remark* . Proposition 7.2.10 implies that any locally convex space $\mathcal{X}$ can be seen as a dual, and in this setting the weak$^*$ topology is precisely the weak topology; in the case where $\mathcal{X}$ is already a dual, this weak$^*$-topology does not agree with the natural weak$^*$-topology. We can say that every locally convex space has a **predual**, i.e. it can be seen as a dual. That said, even when $\mathcal{X}$ is a Banach space this predual need not be a Banach space. For instance, in the situation of Examples 7.2.7, a predual of $c_0$ is $\ell^1(\mathbb{N})$ considered with its weak$^*$-topology. This cannot be a normed space due to Proposition 7.2.9. Let us remark, also, that the predual is not necessarily unique. For instance we know that $c_0^* = c^* = \ell^1(\mathbb{N})$, and by Proposition 7.2.10 we also have $(\ell^\infty(\mathbb{N}), w^*)^* = \ell^1(\mathbb{N})$. Note that while $c_0$ and $c$ are isomorphic (Exercise 5.6.4), they are not isometrically isomorphic (Exercise 5.6.5). And neither of them can be isomorphic to $(\ell^\infty(\mathbb{N}), w^*)$, since in the latter cannot be metrizable by Proposition 7.2.9. The space $\ell^1(\mathbb{N})$ has scores of non-isomorphic preduals, even normed ones.

We will keep talking about preduals in Proposition 7.2.26, Corollary 7.3.8, and later in Section 12.7.

*7.2.12. Remark* . Another consequence of Proposition 7.2.10 is that talk of double duals really belongs to normed spaces. For an arbitrary topological space there is no canonical topology on the dual other than the weak$^*$-topology, and in such case the double dual is the original space, and the double dual contributes nothing.

The most important property of the weak$^*$ topology is the Banach–Alaoglu Theorem:

***7.2.13. Theorem. (Banach–Alaoglu)***

*Let $\mathcal{X}$ be a topological vector space and $V$ an open neighbourhood of 0. Then*

$$K = \{\varphi \in \mathcal{X}^* : \ |\varphi(x)| \leq 1, \ \forall x \in V\}$$

*is weak$^*$-compact.*

*Proof.* Since $V$ is absorbing, for each $x \in \mathcal{X}$ there exists $\gamma(x) \in \mathbb{R}^+$ such that $x \in \gamma(x)V$. So, for every $\varphi \in K$, $|\varphi(x)| \leq \gamma(x)$. Write, for each $x \in \mathcal{X}$, $D_x = \overline{B_{\gamma(x)}(0)}$. Then each $D_x$ is a compact subset of $\mathbb{C}$. By Tychonoff's Theorem (1.8.24), the Cartesian product $D = \prod_{x\in\mathcal{X}} D_x$ is compact in the product topology. We can see $D$ as the set of functions $f : \mathcal{X} \to \mathbb{C}$ such that $|f(x)| \leq \gamma(x)$ for all $x \in \mathcal{X}$, with the pointwise convergence. Note that $K \subset D$.

**Claim 1.** The product topology agrees with $\sigma(\mathcal{X}^*, \mathcal{X})$ on $K$.

**Claim 2.** $K$ is closed in $D$.

The two claims imply the theorem, since Claim 2 implies that $K$ is compact in the product topology, and then Claim 1 implies that $K$ is compact in $\sigma(\mathcal{X}^*, \mathcal{X})$. So we prove both claims.

*Proof of Claim 1.* The product topology on $K$ is given by pointwise convergence. By Proposition 7.2.3, the weak*-topology is also given by pointwise convergence.

*Proof of Claim 2.* Suppose that $\psi$ is in the closure of $K$ in the product topology. Then there exists $\{\varphi_j\} \subset K$ such that $\varphi_j \to \psi$ pointwise. This implies that $\psi$ is linear. Moreover, for any $x \in V$, $|\psi(x)| = |\lim_j \varphi_j(x)| \leq 1$. It remains to see that $\psi$ is continuous. Let $\varepsilon > 0$; then we can consider the open neighbourhood of 0 given by $\varepsilon V$. For any $x \in \varepsilon V$, $x = \varepsilon y$ with $y \in V$, and so $|\psi(x)| = \varepsilon\, |\psi(y)| \leq \varepsilon$. So $\psi$ is continuous, i.e. $\psi \in \mathcal{X}^*$. Thus $\psi \in K$, and $K$ is closed subset of $D$. $\square$

The most typical use of Theorem 7.2.13 occurs when $\mathcal{X}$ is a normed space and $V$ is the open unit ball; in that case what the theorem says is that $\overline{B_1^{\mathcal{X}^*}(0)}$ is weak*-compact. This is the most common situation where Banach–Alaoglu is used.

When we stay with the original topology, the result in Proposition 7.2.10 has to be weakened a bit, but not too much:

***7.2.14. Theorem.***

*Let $\mathcal{X}$ be a normed space. Then $\mathcal{X}$ is weak*-dense in $\mathcal{X}^{**}$. That is,*

$$\overline{\{\hat{x} : \ x \in \mathcal{X}\}}^{\sigma(\mathcal{X}^{**}, \mathcal{X}^*)} = \mathcal{X}^{**}.$$

*Proof.* Suppose that $\psi \in \mathcal{X}^{**}$ but it does not belong to the weak*-closure of $\hat{\mathcal{X}}$. Then we can apply Theorem 5.7.18 (the second geometric version of Hahn–Banach) to the compact set $\{\psi\}$ and the closure of $\hat{\mathcal{X}}$, on the locally convex space $(\mathcal{X}^{**}, \sigma(\mathcal{X}^{**}, \mathcal{X}^*))$. So there is a linear $\varphi : \mathcal{X}^{**} \to \mathbb{C}$, continuous in $\sigma(\mathcal{X}^{**}, \mathcal{X}^*)$ and $c \in \mathbb{R}$ with

$$\operatorname{Re}\varphi(\psi) < c < \operatorname{Re}\varphi(\hat{x}), \quad x \in \mathcal{X}. \tag{7.3}$$

By Proposition 7.2.10, applied to $(\mathcal{X}^{**}, \sigma(\mathcal{X}^{**}, \mathcal{X}^*))$, there exists $\varphi_0 \in \mathcal{X}^*$ such that $\varphi(\psi) = \psi(\varphi_0)$ (for all $\psi$ in $\mathcal{X}^{**}$, but in particular for the one we have fixed). Now $\varphi(\hat{x}) = \hat{x}(\varphi_0) = \varphi_0(x)$, and the inequalities (7.3) become

$$\operatorname{Re}\psi(\varphi_0) < c < \operatorname{Re}\varphi_0(x), \quad x \in \mathcal{X}. \tag{7.4}$$

Among other things, the inequalities imply that $\varphi_0 \neq 0$. Let $x_0 \in \mathcal{X}$ such that $\varphi_0(x_0) = 1$. Let $x_1 = \psi(\varphi_0)x_0 \in \mathcal{X}$. Then

$$\varphi_0(x_1) = \psi(\varphi_0),$$

contradicting (7.4). So $\psi$ cannot exist and thus $\hat{\mathcal{X}}$ is dense. □

*7.2.15. Remark.* In the proof of Theorem 7.2.14, it looks as if one can use the idea with the $x_0$ and $x_1$ already at the level of inequality (7.3). The problem is that we cannot get directly, from $\varphi \neq 0$, that $\varphi(\hat{x}) \neq 0$ for some $x$. For example let $\mathcal{X} = c_0$; then $\mathcal{X}^{**} = \ell^\infty(\mathbb{N})$. If we fix a free ultrafilter $\omega$ and take $\varphi \in \mathcal{X}^{***}$ given by $\varphi(y) = \lim_{n\to\omega} y_n$, then $\varphi$ is a nonzero linear map on $\mathcal{X}^{**}$ such that $\varphi|_{\mathcal{X}} = 0$. Of course, it is not weak*-continuous; with the knowledge of Theorem 7.2.14 we know that if $\varphi$ is weak*-continuous and $\varphi|_{\hat{\mathcal{X}}} = 0$, then $\varphi = 0$, but we could not use that knowledge to prove the theorem! We will see soon, though, that the proof can be achieved with a lot less machinery (cfr. Remark 7.2.19).

As Theorem 7.2.14 shows, $\mathcal{X}$ is always weak*-dense in $\mathcal{X}^{**}$. When the stronger condition $\mathcal{X}^{**} = \mathcal{X}$ holds, we say that $\mathcal{X}$ is **reflexive**; this happens, among other cases, when $\dim \mathcal{X} < \infty$ (Exercise 7.2.9), when $\mathcal{X}$ is a Hilbert space, and more generally when $\mathcal{X} = \ell^p(\mathbb{N})$, $1 < p < \infty$. We will show some characterizations of reflexivity in Proposition 7.2.21. It is important to note that when we write $\mathcal{X}^{**} = \mathcal{X}$ we mean precisely that $\mathcal{X}^{**} = \{\hat{x} : \ x \in \mathcal{X}\}$; it is possible to have a non-reflexive Banach space where $\mathcal{X}^{**} \simeq \mathcal{X}$ isometrically [**Jam51**].

For the remainder of the section we will focus on more specialized results and applications that might be safely skipped on a first approach to the weak*-topology, as they will be needed much later.

***7.2.16. Proposition.***

*Let $\mathcal{X}$ be a reflexive Banach space, and $\mathcal{Y} \subset \mathcal{X}$ a closed subspace. Then $\mathcal{Y}$ is reflexive.*

*Proof.* By hypothesis, the canonical embedding $x \longmapsto \hat{x}$ of $\mathcal{X}$ into $\mathcal{X}^{**}$ is surjective. We already know that it is isometric (see (7.2) on page 440). Similarly, the canonical embedding $y \longmapsto \hat{y}$ from $\mathcal{Y}$ into $\mathcal{Y}^{**}$ is isometric; we want to show it is surjective.

Let $\Psi \in \mathcal{Y}^{**}$. This is a bounded functional $\Psi : \mathcal{Y}^* \to \mathbb{C}$. As $\mathcal{Y} \subset \mathcal{X}$, any $\phi \in \mathcal{X}^*$ restricts to $\mathcal{Y}$, so we may define $\tilde{\Psi} : \mathcal{X}^* \to \mathbb{C}$ by $\tilde{\Psi}(\phi) = \Psi(\phi|_{\mathcal{Y}})$. As the $\Psi$ is bounded and $\|\phi|_{\mathcal{Y}}\| \leq \|\phi\|$, this new map $\tilde{\Psi}$ is bounded, i.e., $\tilde{\Psi} \in \mathcal{X}^{**}$. As $\mathcal{X}$ is reflexive, $\tilde{\Psi} = \hat{x}$ for some $x \in \mathcal{X}$.

Suppose that $x \in \mathcal{X} \backslash \mathcal{Y}$; as $\mathcal{Y}$ is closed, by Hahn–Banach (Corollary 5.7.19) there exists $\varphi \in \mathcal{X}^*$ such that $\varphi(x) = 1$, $\varphi|_{\mathcal{Y}} = 0$. So

$$1 = \varphi(x) = \hat{x}(\varphi) = \tilde{\Psi}(\varphi) = \Psi(\varphi|_{\mathcal{Y}}) = \Psi(0) = 0,$$

a contradiction. Thus $x \in \mathcal{Y}$, and $\Psi = \hat{x}$. $\square$

Goldstine's Theorem below shows that the result in Theorem 7.2.14 can be refined a bit, in the sense that the nets $\{\hat{x}_j\}$ that converge to $\Psi \in \mathcal{X}^{**}$ can be assumed to be bounded. We will use a technical lemma.

**7.2.17. Lemma. (Helly)**

*Let $\mathcal{X}$ be a normed space, $\varphi_1, \ldots, \varphi_n \in \mathcal{X}^*$, and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. The following statements are equivalent:*

*(i) for all $\varepsilon > 0$ there exists $x \in B_1^{\mathcal{X}}(0)$ with $|\varphi_k(x) - \alpha_k| < \varepsilon$, $k = 1, \ldots, n$;*

*(ii) for all $\beta_1, \ldots, \beta_n \in \mathbb{C}$,*

$$\Big|\sum_{k=1}^n \beta_k \alpha_k\Big| \leq \Big\|\sum_{k=1}^n \beta_k \varphi_k\Big\|. \tag{7.5}$$

*Proof.* (i) $\Longrightarrow$ (ii) Let $\beta = \sum_k |\beta_k|$. Then, since $\|x\| \leq 1$,

$$\Big|\sum_{k=1}^n \beta_k \alpha_k\Big| \leq \Big|\sum_{k=1}^n \beta_k \varphi_k(x)\Big| + \Big|\sum_{k=1}^n \beta_k (\varphi_k(x) - \alpha_k)\Big| \leq \Big\|\sum_{k=1}^n \beta_k \varphi_k\Big\| + \varepsilon\,\beta.$$

As this works for any $\varepsilon > 0$, (7.5) holds.

(ii) $\Longrightarrow$ (i) Let $\alpha = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix}^\top \in \mathbb{C}^n$ and define $\gamma : \mathcal{X} \to \mathbb{C}^n$ by

$$\gamma(x) = \begin{bmatrix} \varphi_1(x) & \cdots & \varphi_n(x) \end{bmatrix}^\top.$$

As $\gamma$ is linear and bounded, the set $\gamma(B_1^{\mathcal{X}}(0))$ is convex and bounded; so $\overline{\gamma(B_1^{\mathcal{X}}(0))}$ is convex (Exercise 5.4.6) and compact.

If we assume that (i) is false, then $\alpha \notin \overline{\gamma(B_1^{\mathcal{X}}(0))}$ (Exercise 7.2.10). So we have a point and a convex compact set in $\mathbb{C}^n$, that are disjoint. Then Hahn–Banach (Theorem 5.7.18), combined with the Riesz Representation Theorem (Theorem 4.5.4), give us $c \in \mathbb{R}$ and $\beta \in \mathbb{C}^n$ such that

$$\operatorname{Re}\langle \gamma(x), \beta\rangle < c < \operatorname{Re}\langle \alpha, \beta\rangle, \qquad x \in B_1^{\mathcal{X}}(0).$$

Because $\overline{B_1^{\mathcal{X}}(0)}$ is balanced, we get

$$|\langle \gamma(x), \beta\rangle| < c < \operatorname{Re}\langle \alpha, \beta\rangle \leq |\langle \alpha, \beta\rangle|, \qquad x \in B_1^{\mathcal{X}}(0).$$

If we expand the above, writing $\overline{\beta_k}$ for the components of $\beta$, we have

$$\Big|\sum_{k=1}^{n} \beta_k \varphi_k(x)\Big| < c < \Big|\sum_{k=1}^{n} \beta_k \alpha_k\Big|.$$

As this can be done for all $x \in B_1^{\mathcal{X}}(0)$,

$$\Big\|\sum_{k=1}^{n} \beta_k \varphi_k\Big\| \le c < \Big|\sum_{k=1}^{n} \beta_k \alpha_k\Big|,$$

contradicting (ii). □

**7.2.18. Theorem. (Goldstine)**

*Let $\mathcal{X}$ be a normed space. Then the closed unit ball of $\hat{\mathcal{X}}$ is weak*-dense in the closed unit ball of $\mathcal{X}^{**}$. That is,*

$$\overline{B_1^{\hat{\mathcal{X}}}(0)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)} = \overline{B_1^{\mathcal{X}^{**}}(0)}.$$

*Proof.* Let $\psi \in \mathcal{X}^{**}$ with $\|\psi\| \le 1$, and fix a weak*-neighbourhood

$$\psi + N(\varphi_1, \dots, \varphi_n; 1)$$

of $\psi$. Let $\alpha_k = \psi(\varphi_k)$, $k = 1, \dots, n$. Given $\beta_1, \dots, \beta_n \in \mathbb{C}$,

$$\Big|\sum_{k=1}^{n} \beta_k \alpha_k\Big| = \Big|\sum_{k=1}^{n} \beta_k \psi(\varphi_k)\Big| = \Big|\psi\Big(\sum_{k=1}^{n} \beta_k \varphi_k\Big)\Big| \le \Big\|\sum_{k=1}^{n} \beta_k \varphi_k\Big\|.$$

By Lemma 7.2.17 there exists $x \in B_1^{\mathcal{X}}(0)$ such that $|\varphi_k(x) - \alpha_k| < 1$ for all $k$. If we rewrite this as $|\hat{x}(\varphi_k) - \alpha_k| < 1$ for all $k$, this shows that $\hat{x} \in \psi + N(\varphi_1, \dots, \varphi_n; 1)$. This way the net $\{\hat{x}_N\}$, indexed by the open weak*-neighbourhoods of $\psi$ ordered by inclusion, converges to $\psi$. □

*7.2.19. Remark*. The proof of Theorem 7.2.18 gives us a different proof of Theorem 7.2.14. It uses a lot less machinery, because Lemma 7.2.17 only uses the Hahn–Banach Theorem and the Riesz Representation Theorem in $\mathbb{C}^n$.

**7.2.20. Corollary.**

*Let $\mathcal{X}$ be a normed space. The following statements are equivalent:*

*(i) when considered with the weak*-topology, the closed unit ball of $\mathcal{X}^*$ is a compact metrizable space;*

*(ii)* $\mathcal{X}$ *is separable.*

*Proof.* (i) $\Longrightarrow$ (ii) Suppose that $\overline{B_1^{\mathcal{X}^*}(0)}$ is $\sigma(\mathcal{X}^*, \mathcal{X})$-metrizable. As we will see in Proposition 7.4.24, this implies that the space $C\big(\overline{B_1^{\mathcal{X}^*}(0)}\big)$ is separable. Let $\Gamma : \mathcal{X} \to C\big(\overline{B_1^{\mathcal{X}^*}(0)}\big)$ be $\Gamma(x) = \hat{x}$. We have

$$\|\Gamma(x)\| = \max\{|\hat{x}(\varphi)| : \ \varphi \in B_1^{\mathcal{X}^*}(0)\} = \max\{|\varphi(x)| : \ \varphi \in B_1^{\mathcal{X}^*}(0)\} = \|x\|,$$

so $\Gamma$ is an isometry. As $\Gamma(\mathcal{X})$ is a subspace of a separable metric space it is separable (Proposition 1.8.5), and hence $\mathcal{X}$ is separable.

(ii) $\Longrightarrow$ (i) We already know, by Theorem 7.2.13, that $\overline{B_1^{\mathcal{X}^*}(0)}$ is weak*-compact. Let $\{x_n\} \subset B_1^{\mathcal{X}}(0)$ be a dense sequence. Define

$$d(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n} |\varphi(x_n) - \psi(x_n)|, \qquad \varphi, \psi \in \overline{B_1^{\mathcal{X}^*}(0)}.$$

The series converges, since $|\varphi(x_n) - \psi(x_n)| \le \|x_n\| \left(\|\varphi\| + \|\psi\|\right) \le 2$. It is also translation invariant, symmetric and satisfies the triangle inequality, so it is a translation invariant metric. If $\varphi_j \xrightarrow{\text{weak}^*} 0$, fix $\varepsilon > 0$ and choose $n_0$ such that $\sum_{n>n_0} 2^{-n} < \varepsilon/4$. Choose also $j_0$ such that $|\varphi_j(x_n)| < \varepsilon/2$, if $j > j_0$, $n = 1, \ldots, n_0$. Then, for $j > j_0$,

$$\begin{aligned} d(\varphi_j, 0) &= \sum_{n=1}^{n_0} 2^{-n} |\varphi(x_n) - \psi(x_n)| + \sum_{n>n_0} 2^{-n} |\varphi(x_n) - \psi(x_n)| \\ &< \sum_{n=1}^{n_0} 2^{-n} \frac{\varepsilon}{2} + 2\,\frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

It follows that $d(\varphi_j, 0) \to 0$. Conversely, suppose that $d(\varphi_j, 0) \to 0$. Fix $x \in \overline{B_1^{\mathcal{X}}(0)}$ and $\varepsilon > 0$. There exists $n$ such that $\|x - x_n\| < \varepsilon$. Then

$$|\varphi_j(x)| \le |\varphi_j(x - x_n)| + |\varphi_j(x_n)| \le \|x - x_n\| + |\varphi_j(x_n)| \le \varepsilon + 2^n\, d(\varphi_j, 0).$$

Thus $\limsup_j |\varphi_j(x)| \le \varepsilon$. By the Limsup Routine, $\lim_j \varphi_j(x) = 0$. $\square$

A consequence of Goldstine's Theorem is the following characterization of reflexivity. We note that since the double dual of a normed space is always Banach, a reflexive normed space is necessarily Banach.

***7.2.21. Proposition.***

*Let $\mathcal{X}$ be a normed space. The following statements are equivalent.*

*(i)* $\mathcal{X}$ *is reflexive;*

*(ii)* $\sigma(\mathcal{X}^*, \mathcal{X}) = \sigma(\mathcal{X}^*, \mathcal{X}^{**})$*;*

*(iii)* $\overline{B_1^{\mathcal{X}}(0)}$ *is weakly compact.*

*Proof.* (i) $\implies$ (ii) By hypothesis every $\Gamma \in \mathcal{X}^{**}$ is of the form $\hat{x}$ for some $x \in \mathcal{X}$. So $\varphi_j \xrightarrow{\text{weak}} 0 \iff \Gamma(\varphi_j) \to 0,\ \Gamma \in \mathcal{X}^{**} \iff \hat{x}(\varphi_j) \to 0,\ x \in \mathcal{X} \iff \varphi_j(x) \to 0,\ x \in \mathcal{X} \iff \varphi_j \xrightarrow{\text{weak}^*} 0$.

(ii) $\implies$ (iii) By Banach–Alaoglu (Theorem 7.2.13), the closed unit ball $\overline{B_1^{\mathcal{X}^{**}}(0)}$ is weak*-compact. As the two topologies agree, $\overline{B_1^{\mathcal{X}}(0)}$ is weakly compact.

(iii) $\implies$ (i) By Goldstine's Theorem (Theorem 7.2.18),

$$\overline{B_1^{\mathcal{X}^{**}}(0)} = \overline{B_1^{\mathcal{X}}(0)}^{\text{weak}^*} = \overline{B_1^{\mathcal{X}}(0)}^{\text{weak}} = \overline{B_1^{\mathcal{X}}(0)},$$

the second to last equality because since $\overline{B_1^{\mathcal{X}}(0)}$ is weakly compact, it is weak*-compact as a subset of $\overline{B_1^{\mathcal{X}^{**}}(0)}$. And the last equality can be checked directly; it also follows from Theorem 7.1.16. Then $\mathcal{X}^{**} = \mathcal{X}$, and $\mathcal{X}$ is reflexive. $\square$

We saw in Examples 7.2.7 an example of a dual (of the non-reflexive space $c_0$) where the weak and weak*-topologies do not agree. Now we know, via Proposition 7.2.21, that the dual of any non-reflexive Banach space will provide an example.

### *7.2.22. Corollary.*

*Let* $\mathcal{X}$ *be a Banach space. Then* $\mathcal{X}$ *is reflexive if and only if* $\mathcal{X}^*$ *is reflexive.*

*Proof.* Assume first that $\mathcal{X}$ is reflexive. Then $\mathcal{X} = \mathcal{X}^{**}$. By Proposition 7.2.21 $\sigma(\mathcal{X}^*, \mathcal{X}) = \sigma(\mathcal{X}^*, \mathcal{X}^{**})$. Then $\overline{B_1^{\mathcal{X}^*}(0)}$ is weak compact by Banach–Alaoglu, and so $\mathcal{X}^*$ is reflexive again by Proposition 7.2.21.

Conversely, if $\mathcal{X}^*$ is reflexive then $\mathcal{X}^{**}$ is reflexive by the first part of the proof. By Proposition 7.2.21, $\overline{B_1^{\mathcal{X}^{**}}(0)}$ is weak compact. The unit ball $\overline{B_1^{\mathcal{X}}(0)}$ is a weakly closed subset of $\overline{B_1^{\mathcal{X}^{**}}(0)}$ (by Theorem 7.1.16) and so it is weakly compact (careful: this is the weak topology in $\mathcal{X}^{**}$). But the weak topology $\sigma(\mathcal{X}, \mathcal{X}^*)$ of $\mathcal{X}$ is weaker than $\sigma(\mathcal{X}^{**}, \mathcal{X}^{***})$. So $\overline{B_1^{\mathcal{X}}(0)}$ is weak*-compact in $\mathcal{X}^{**}$, which is precisely $\overline{B_1^{\mathcal{X}}(0)}$ being compact in its weak topology $\sigma(\mathcal{X}, \mathcal{X}^*)$. Then $\mathcal{X}$ is reflexive by Proposition 7.2.21. $\square$

Recall the definition of the $\ell^p$ direct sum of Banach spaces (Definition 5.3.8).

**7.2.23. Proposition.**

*Let $\{\mathcal{X}_n\}$ be a sequence of reflexive Banach spaces, $1 < p < \infty$, and let $\mathcal{X} = \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}$ be the $\ell^p$-direct sum. Then $\mathcal{X}$ is reflexive.*

*Proof.* Using Proposition 5.6.5 twice, we have

$$\Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}^{**} \simeq \Big(\bigoplus_n \mathcal{X}_n^*\Big)_{\ell^q}^{*} \simeq \Big(\bigoplus_n \mathcal{X}_n^{**}\Big)_{\ell^p} = \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}. \tag{7.6}$$

We are not done, because we do not need "some" isomorphism but the canonical one. If $\bigoplus_n x_n \in \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}$, the isomorphisms above map

$$\widehat{\bigoplus_n x_n} \longmapsto \bigoplus_n \hat{x}_n,$$

so it is indeed the canonical isomorphism and this shows that $\mathcal{X}$ is reflexive. To check this, in the first isomorphism we are considering the map $(\Gamma'_*)^*$, where $\Gamma'_* : \Big(\bigoplus_n \mathcal{X}_n^*\Big)_{\ell^q} \to \Big(\bigoplus_n \mathcal{X}_n\Big)_{\ell^p}^*$ is the reverse map in Proposition 5.6.5 applied to the space $\Big(\bigoplus_n \mathcal{X}_n^*\Big)_{\ell^q}$. So

$$\begin{aligned}(\Gamma'_*)^*\Big(\widehat{\bigoplus_n x_n}\Big)\Big(\bigoplus_n \varphi_n\Big) &= \Big(\widehat{\bigoplus_n x_n}\Big)\Big(\Gamma'_*\Big(\bigoplus_n \varphi_n\Big)\Big) = \Gamma'_*\Big(\bigoplus_n \varphi_n\Big)\Big(\bigoplus_n x_n\Big) \\ &= \sum_n \varphi_n(x_n) = \Big(\bigoplus_n \hat{x}_n\Big)\Big(\bigoplus_n \varphi_n\Big).\end{aligned}$$

The second isomorphism in (7.6) is now trivial, because we are applying $\Gamma$ to a functional that is already a direct sum. $\square$

We record here an unexpected consequence of the above results, that we was shown in the proof of Corollary 7.2.20.

**7.2.24. Proposition.**

*Let $\mathcal{X}$ be a separable normed space. Then there exists a metrizable compact Hausdorff space $K$ such that $\mathcal{X}$ embeds isometrically into $C(K)$.*

The result in Proposition 7.2.24 should not be seen as an indication that there is no variety of Banach spaces. With more sophisticated ideas, though, one can take $K$ to be $[0,1]$ (Theorem 7.6.6).

If $\mathcal{X}$ is a Banach space and $K \subset \mathcal{X}$ is convex, then $K$ is weakly closed if and only if $K \cap \overline{B_r(0)}$ is weakly closed for all $r > 0$ (Exercise 7.1.16). Note that it is not necessary to specify whether the closure of the balls is the weak or the norm one; the key to the proof is Theorem 7.1.16. We cannot expect the argument to work for the weak*-topology, though, since the weak and weak* closures of a convex set do not generally agree (cfr. Examples 7.2.7). Remarkably, the result still holds for the weak*-topology.

***7.2.25. Theorem. (Krein–Šmulian)***

*Let $\mathcal{X}$ be a Banach space, and $K \subset \mathcal{X}^*$ be convex. Then $K$ is weak*-closed if and only if $K \cap \overline{B_r^{\mathcal{X}^*}(0)}$ is weak*-closed for all $r > 0$.*

*Proof.* If $K$ is weak*-closed then the intersection is weak*-closed, for $\overline{B_r^{\mathcal{X}^*}(0)}$ is weak*-closed by Banach–Alaouglu.

For the converse, we assume that $K \cap \overline{B_r^{\mathcal{X}^*}(0)}$ is weak*-closed for all $r > 0$. First we show that $K$ is norm-closed. Let $\varphi \in \overline{K}$. So there exists $\{\varphi_n\} \subset K$ with $\|\varphi_n - \varphi\| \to 0$. Then there exists $r > 0$ such that $\|\varphi_n\| \leq r$ for all $n$. So the sequence $\{\varphi_n\}$ converges in norm, hence weak*, to $\varphi$ in the weak*-closed set $K \cap \overline{B_r^{\mathcal{X}^*}(0)}$; thus $\varphi \in K$ and $K$ is norm closed. Next suppose that $\varphi \in \overline{K}^{w^*} \setminus K$. As $K$ is norm-closed, there exists $s > 0$ such that $\overline{B_s^{\mathcal{X}^*}(\varphi)} \cap K = \varnothing$. This implies that $\overline{B_1^{\mathcal{X}^*}(0)} \cap \frac{1}{s}(K - \varphi) = \varnothing$; indeed, if $\psi \in \frac{1}{s}(K - \varphi)$ and $\|\psi\| \leq 1$, then $\varphi + s\psi \in K \cap \overline{B_s(\varphi)}$, a contradiction. For $n$ big enough we have $[\frac{1}{s}(K - \varphi)] \cap \overline{B_n^{\mathcal{X}^*}(0)} \neq \varnothing$ (namely, any $n \geq \frac{1}{s}\|\psi - \varphi\|$ for any $\psi \in K$ works), and $[\frac{1}{s}(K - \varphi)] \cap \overline{B_n^{\mathcal{X}^*}(0)}$ is weak*-closed; for if $\{\psi_j\} \subset [\frac{1}{s}(K - \varphi)] \cap \overline{B_n^{\mathcal{X}^*}(0)}$ and $\psi_j \xrightarrow{\text{weak}^*} \psi$, then $s\psi_j + \varphi \in K \cap \overline{B_{sn+\|\varphi\|}^{\mathcal{X}^*}(0)}$ since it is weak*-closed by hypothesis, so $\psi \in \frac{1}{s}(K - \varphi)$ and hence $\psi \in [\frac{1}{s}(K - \varphi)] \cap \overline{B_n^{\mathcal{X}^*}(0)}$ (Exercise 7.2.6).

By Hahn–Banach (Theorem 5.7.18), applied in the locally convex space $(\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$, to the weak*-compact set $\overline{B_1^{\mathcal{X}^*}(0)}$ and the disjoint weak*-closed convex set $[\frac{1}{s}(K - \varphi)] \cap \overline{B_n^{\mathcal{X}^*}(0)}$ we get by Proposition 7.2.10 that there exists $x_n \in \mathcal{X}$ such that

$$\operatorname{Re}\psi(x_n) < 1 \leq \operatorname{Re}\eta(x_n), \qquad \psi \in \overline{B_1^{\mathcal{X}^*}(0)},\ \eta \in \frac{1}{s}(K - \varphi)\overline{B_n^{\mathcal{X}^*}(0)}. \tag{7.7}$$

For a fixed $\psi \in B_1^{\mathcal{X}^*}(0)$, write $\psi(x_n) = \lambda\,|\psi(x_n)|$, $\lambda \in \mathbb{T}$. Then

$$|\psi(x_n)| = \bar{\lambda}\,\psi(x_n) = \operatorname{Re}(\bar{\lambda}\,\psi)(x_n) < 1$$

since $\bar{\lambda}\psi \in B_1^{\mathcal{X}^*}(0)$. This shows that $|\psi(x_n)| < 1$ for all $\psi \in B_1^{\mathcal{X}^*}(0)$, and therefore $\|x_n\| \leq 1$ by Corollary 5.7.7. After repeating this for all (big enough) $n$, we have $\{x_n\} \subset \overline{B_1^{\mathcal{X}^*}(0)}$, and by Banach–Alaoglu there exists a weak*-cluster point $x$. From (7.7),

$$\operatorname{Re}\psi(x) \leq 1 \leq \operatorname{Re}\eta(x), \qquad \psi \in \overline{B_1^{\mathcal{X}^*}(0)},\ \eta \in \frac{1}{s}\,(K-\varphi). \tag{7.8}$$

Now choose $\{\varphi_j\} \subset K$ with $\varphi_j \xrightarrow{\text{weak}^*} \varphi$. Since $\frac{1}{s}(\varphi_j - \varphi) \in \frac{1}{s}\,(K-\varphi)$, we obtain from (7.8) that $\operatorname{Re}\left(\frac{1}{s}\,(\varphi_j(x) - \varphi(x)\right) \geq 1$, which we can rewrite as

$$\operatorname{Re}\big(\varphi_j(x) - \varphi(x)\big) \geq s.$$

Since $\varphi_j(x) \to \varphi(x)$, this implies that $0 \geq s$, a contradiction. It follows that $\varphi$ cannot exist and thus $\overline{K}^{w^*} = K$. ☐

The reader might note that the completeness of $\mathcal{X}$ is never mentioned in the proof. But it is actually used for the existence of the cluster point $x$.

We will see in a bit that not every Banach space is the dual of another Banach space; concretely, neither $L^1[0,1]$ nor $c_0$ are duals (Proposition 7.5.13). This raises the question of which Banach spaces are duals and which aren't. Suppose that $\mathcal{X} = \mathcal{Y}^*$ for some normed space $\mathcal{Y}$. The embedding from Corollary 7.2.5 tells us that we can see $\mathcal{Y}$ as a subspace of $\mathcal{X}^*$. So, any predual of $\mathcal{X}$ can somehow be seen as a subspace of $\mathcal{X}^*$. And these can be characterized:

***7.2.26. Proposition.***

*Let $\mathcal{X}$ be a Banach space, and $\mathcal{Y} \subset \mathcal{X}^*$ a closed subspace. The following statements are equivalent:*

*(i) $\mathcal{Y}$ is a predual for $\mathcal{X}$;*

*(ii) (a) for all $x \in \mathcal{X}$, $\|x\| = \sup\{|\varphi(x)| :\ \varphi \in \mathcal{Y},\ \|\varphi\| = 1\}$,*
*(b) $\overline{B_1^{\mathcal{X}}(0)}$ is compact in the $\sigma(\mathcal{X},\mathcal{Y})$ topology.*

*Proof.* (i) $\Longrightarrow$ (ii) This is Proposition 5.5.6 and Theorem 7.2.13.

(ii) $\Longrightarrow$ (i) We define $\Gamma : \mathcal{X} \to \mathcal{Y}^*$ via the usual duality $\hat{x}(\varphi) = \varphi(x)$; that is, $\Gamma(x) = \hat{x}$. We have

$$\|\Gamma(x)\| = \sup\{|\hat{x}(\varphi)| :\ \varphi \in B_1^{\mathcal{Y}}(0)\} = \sup\{|\varphi(x)| :\ \varphi \in B_1^{\mathcal{Y}}(0)\} = \|x\|,$$

the last inequality by the first hypothesis. So $\Gamma$ is a linear isometry. Due to the definition of $\hat{x}$, $\Gamma$ is also continuous from $(\mathcal{X}, \sigma(\mathcal{X},\mathcal{Y}))$ to $(\mathcal{Y}^*, \sigma(\mathcal{Y}^*,\mathcal{Y}))$; indeed, if $x_j \to x$ in $\sigma(\mathcal{X},\mathcal{Y})$, this means that $\varphi(x_j) \to \varphi(x)$ for all $\varphi \in \mathcal{Y}$, and this is exactly that $\hat{x}_j(\varphi) \to \hat{x}(\varphi)$. By the second hypothesis, this means that $\Gamma(\overline{B_1^{\mathcal{X}}(0)})$ is $\sigma(\mathcal{Y}^*,\mathcal{Y})$-compact in $\mathcal{Y}^*$. Thus $\Gamma$ is a homeomorphism (from $(\mathcal{X},\sigma(\mathcal{X},\mathcal{Y}))$ to $(\mathcal{Y}^*,\sigma(\mathcal{Y}^*,\mathcal{Y}))$ onto its image.

What remains is to show that $\Gamma$ is surjective. Because $\Gamma$ is isometric, for any $r>0$

$$\Gamma(\overline{B_r^{\mathcal{X}}(0)}) = \overline{\Gamma(B_r^{\mathcal{X}}(0))} = \overline{\Gamma(\mathcal{X})\cap B_r^{\mathcal{Y}^*}(0)} = \Gamma(\mathcal{X})\cap \overline{B_r^{\mathcal{Y}^*}(0)}.$$

We saw in the previous paragraph that the set above is weak$^*$-compact in $\mathcal{Y}^*$. By Krein–Šmulian, $\Gamma(\mathcal{X})$ is weak$^*$-closed. If $\Phi\in\mathcal{Y}^*\setminus\Gamma(\mathcal{X})$, by Corollary 5.7.19, applied on $(\mathcal{Y}^*,\sigma(\mathcal{Y}^*,\mathcal{Y}))$, there exists $\gamma$ in the double dual with $\gamma(\Phi)=1$, and $\gamma(\Gamma(x))=0$ for all $x\in\mathcal{X}$. By Proposition 7.2.10, this $\gamma$ is implemented by evaluation at a certain $\varphi\in\mathcal{Y}$. Then we would have

$$0=\gamma(\Gamma(x)) = (\Gamma(x))(\varphi) = \hat{x}(\varphi) = \varphi(x).$$

But this is impossible unless $x=0$, for the norm of $x$ can be obtained over $\mathcal{Y}$ by hypothesis. It follows that $\Gamma(\mathcal{X})=\mathcal{Y}^*$. □

Let us now fulfill, after a short lemma, a promise from Section 7.1 (made on page 442).

**7.2.27. Lemma.**

*Let $\mathcal{X}$ be a normed space and $F\subset\mathcal{X}^*$ a finite-dimensional subspace. Then there exists $n\in\mathbb{N}$ and $x_1,\dots,x_n\in\mathcal{X}$, with $\|x_j\|=1$ for all $j$, such that*

$$\max\left\{|\psi(x_j)| :\ 1\le j\le n\right\} \ge \frac{1}{2}\,\|\psi\|, \qquad \psi\in F. \tag{7.9}$$

*Proof.* Fix $\psi\in F$ with $\|\psi\|=1$. Since the unit sphere of $F$ is compact (being a closed subset of a compact set, see Theorem 5.2.9), it can be covered by finitely many balls of radius $1/4$ (Lemma 1.8.25); that is, there exist $\varphi_1,\dots,\varphi_n\in F$, with $\|\varphi_j\|=1$ for all $j$ and such that $\|\psi-\varphi_{j'}\|<\frac14$ for some $j$'. For each $j$ there exists $x_j\in\mathcal{X}^*$ with $\|x_j\|=1$ and $|\varphi_j(x_j)|>\frac34$. Then

$$|\psi(x_{j'})| \ge |\varphi_{j'}(x_{j'})| - \|\psi-\varphi_{j'}\| > \frac34-\frac14=\frac12.$$

Then (7.9) follows. □

**7.2.28. Theorem. (Eberlein–Šmulian)**

*Let $\mathcal{X}$ be a Banach space and $K\subset\mathcal{X}$. The following statements are equivalent:*

*(i) $K$ is relatively weakly compact;*

*(ii) every sequence in $K$ has a weakly convergent subsequence in $\mathcal{X}$;*

*(iii) every sequence in $K$ has a weak cluster point in $\mathcal{X}$ (that is, every sequence has a convergent subnet).*

*Proof.* We follow the proof from [Whi67]. An alternative proof using basic sequences (that are studied in Section 9.8) was given by Pełczyński[Pel64].

(i) $\implies$ (ii) We assume that $K$ is weakly compact, as otherwise we can work on $\overline{K}^{\sigma(\mathcal{X},\mathcal{X}^*)}$. Let $\{x_n\} \subset K$ be a sequence. We define $\mathcal{X}_0 = \overline{\operatorname{span}}^{\|\cdot\|}\{x_n\}$, and $K_0 = K \cap \mathcal{X}_0$. We have $\{x_n\} \subset K_0$ and $K_0$ is weakly compact (note that, being a norm-closed subspace , $\mathcal{X}_0$ is weakly closed by Theorem 7.1.16). The reason we do this is that $\mathcal{X}_0$ is a separable Banach space, and by Hahn–Banach the weak topology on $\mathcal{X}_0$ is the restriction of the weak topology on $\mathcal{X}$. Given $\{d_n\} \subset \mathcal{X}_0$ dense, by Hahn–Banach (Corollary 5.7.7) there exists, for each $n \in \mathbb{N}$, $\varphi_n \in \mathcal{X}_0^*$ with $\|\varphi_n\| = 1$ and $\varphi_n(d_n) = \|d_n\|$. If we define

$$\rho(x,y) = \sum_{k=1}^{\infty} 2^{-k}\,|\varphi_k(x-y)|, \qquad\qquad x,y \in \mathcal{X},$$

then $\rho$ is a metric: the symmetry and the triangle inequality are immediate; and if $\rho(x,y) = 0$, this means that $\varphi_n(x-y) = 0$ for all $n$, which implies that $x - y = 0$. Indeed, if $x - y \neq 0$ we can take $d_k$ such that $\|d_k - (x-y)\| < \|x-y\|/2$, and we have $\varphi_k(d_k-(x-y)) = \varphi_k(d_k) = \|d_k\|$; from the inequality we get $\|x-y\|/2 \leq \|d_k\|$, and then

$$\frac{1}{2}\,\|x-y\| \leq \|d_k\| = \varphi_k(d_k - (x-y)) \leq \|d_k - (x-y)\|,$$

a contradiction. So $x - y = 0$ and $\rho$ is a translation-invariant metric. We claim that this metric gives the relative weak topology on any weakly compact subset of $\mathcal{X}_0$. To see this, note first that $K$ is bounded by Proposition 7.1.12, so there exists $c > 0$ with $\|x\| < c$ for all $x \in K$. Consider the map $T : (K_0, \sigma(\mathcal{X},\mathcal{X}^*)) \to (K_0, \rho)$ where $T$ is the identity. If we show that $T$ is weakly continuous, since its domain is compact it becomes a homeomorphism onto its image (Exercise 1.8.38), and so the two topologies are the same on $K_0$. To check the continuity, fix $\varepsilon > 0$ and suppose that we have a net $\{z_j\} \subset K_0$ with $z_j \xrightarrow{\text{weak}} z$. From

$$\sum_{k=m+1}^{\infty} 2^{-k}\,|\varphi_k(z_j - z)| \leq 2c \sum_{k=m+1}^{\infty} 2^{-k} = c\,2^{-m+1}, \tag{7.10}$$

there exists $m$ such that the series in (7.10) is less than $\varepsilon/2$. If we choose $j_0$ so that for all $j \geq j_0$ and $k \leq m$ we have $|\varphi_n(z_j - z)| < \frac{\varepsilon}{2}$, then

$$d(z_j, z) = \sum_{k=1}^{m} 2^{-k} |\varphi_k(z_j - z)| + \sum_{k=m+1}^{\infty} 2^{-k} |\varphi_k(z_j - z)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

showing that $T$ is continuous. We have shown that the weak topology is metrizable on $K_0$, and so $\{x_n\}$ admits a convergent subsequence (since it is a bounded sequence in a compact set in a metric space).

(ii) $\implies$ (iii) Trivial.

(iii) $\implies$ (i) For each $\varphi \in \mathcal{X}^*$, as it is weakly continuous the set $\varphi(K) \subset \mathbb{C}$ inherits the property of having a cluster point for each sequence (given a sequence $\{g(x_n)\} \subset \varphi(K)$, there exists a weakly convergent subnet $\{x_{n_j}\}$ and by the weak continuity of $\varphi$ the net $\{\varphi(x_{n_j})\}$ converges in $\mathbb{C}$). This in particular means that $\varphi(K)$ cannot have unbounded sequences, and so $\varphi(K)$ is bounded. We can write this as

$$\sup\{\hat{k}(\varphi) : \; k \in K\} < \infty, \qquad\qquad \varphi \in \mathcal{X}^*.$$

By the Uniform Boundedness Principle (Theorem 6.3.16) $\sup\{\|\hat{k}\| : \; k \in K\} < \infty$, and since $\|\hat{k}\| = \|k\|$ (Exercise 7.2.1) we have shown that $K$ is bounded. As before we denote by $\iota : \mathcal{X} \to \mathcal{X}^{**}$ the canonical embedding. We have that $\iota(K)$ is bounded (since $K$ is bounded and $\iota$ is an isometry), and so $\overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)}$ is weak*-compact by Banach–Alaoglu (Theorem 7.2.13). If we show that $\overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)} \subset \iota(\mathcal{X})$, then using that $\iota$ is a homeomorphism from $\mathcal{X}$ with the weak topology to $\iota(\mathcal{X})$ with the weak*-topology (Corollary 7.2.5) we get

$$K \subset \iota^{-1}(\overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)})$$

and hence $K$ is relatively weakly compact since it lies inside a weakly compact set. So the rest of the proof consists of showing the inclusion $\overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)} \subset \iota(\mathcal{X})$.

We will inductively construct a sequence where the hypothesis can be used. Fix $\psi \in \overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)}$. We want to show that $\psi = \iota(x)$ for some $x \in \mathcal{X}$. Let $\varphi_1 \in \mathcal{X}^*$ with $\|\varphi_1\| = 1$. From $\psi \in \overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)}$ we get that there exists $k_1 \in K$ with $|(\psi - \iota(k_1))\varphi_1| < 1$ (recall the form of the weak* basic neighbourhoods, Remark 7.2.4). Applying Lemma 7.2.27 to $\operatorname{span}\{\psi, \psi - \iota(k_1)\} \subset \mathcal{X}^{**}$ we get $\varphi_2, \ldots, \varphi_{n(2)} \in \mathcal{X}^*$, with $\|\varphi_j\| = 1$ for all $j$ and

$$\max\left\{|\psi'(\varphi_j)| : \; 2 \leq j \leq n(2)\right\} \geq \frac{1}{2}\,\|\psi'\|, \qquad\qquad \psi' \in \operatorname{span}\{\psi, \psi - \iota(k_1)\}.$$

For induction, suppose we have $k_1, \ldots, k_m \in K$ and $\varphi_1, \ldots, \varphi_{n(m)} \in \mathcal{X}^*$ with $\|\varphi_\ell\| = 1$ for all $\ell$ and with $|(\psi - \iota(k_j))\varphi_\ell| < \frac{1}{m}$ and numbers $1 = n(1) <$

$n(2) < \cdots < n(m) \in \mathbb{N}$ with

$$\max\left\{|\psi'(\varphi_\ell)| : \ n(m-1) < \ell \leq n(m)\right\} \geq \frac{1}{2}\left\|\psi'\right\| \tag{7.11}$$

for all $\psi' \in \operatorname{span}\{\psi, \psi - \iota(k_1), \ldots, \psi - \iota(k_m)\}$. Using that $\psi \in \overline{\iota(K)}^{\sigma(\mathcal{X}^{**},\mathcal{X}^*)}$, there exists $k_{m+1} \in K$ with $|(\psi - \iota(k_{m+1}))\varphi_\ell| < \frac{1}{m+1}$ for $\ell = 1, \ldots, n(m)$. Applying Lemma 7.2.27 to $F = \operatorname{span}\{\psi, \psi - \iota(k_1), \ldots, \psi - \iota(k_{m+1})\}$ we get $\varphi_{n(m)+1}, \ldots, \varphi_{n(m+1)} \in \mathcal{X}^*$, of norm one, such that

$$\max\left\{|\psi'(\varphi_\ell)| : \ n(m) < \ell \leq n(m+1)\right\} \geq \frac{1}{2}\left\|\psi'\right\|$$

for all $\psi' \in \operatorname{span}\{\psi, \psi - \iota(k_1), \ldots, \psi - \iota(k_{m+1})\}$, and the induction is complete.

By hypothesis the sequence $\{k_m\} \subset K$ has a weak cluster point $x \in \mathcal{X}$. By Theorem 7.1.16 the subspace $\overline{\operatorname{span}}^{\|\cdot\|}\{k_m\}$ is weakly closed, and so $x \in \overline{\operatorname{span}}^{\|\cdot\|}\{k_m\}$. As $\iota$ is an isometry, we get that $\psi - \iota(x) \in \overline{\operatorname{span}}^{\|\cdot\|}\{\psi, \psi - \iota(k_1), \psi - \iota(k_2), \ldots\}$. Since (7.11) holds in $\operatorname{span}\{\psi, \psi - \iota(k_1), \ldots, \psi - \iota(k_m)\}$, we get by taking limits that

$$\sup\left\{|(\psi - \iota(x))(\varphi_\ell)| : \ \ell \in \mathbb{N}\right\} \geq \frac{1}{2}\left\|\psi - \iota(x)\right\|; \tag{7.12}$$

indeed, if $\varepsilon > 0$ is given then there exists $m$ and $\psi' \in \operatorname{span}\{\psi, \psi - \iota(k_1), \psi - \iota(k_2), \ldots\}$ with $\|\psi - \iota(x) - \psi'\| < \varepsilon$. There is also $\ell$ such that $|\psi'(\varphi_\ell)| \geq \frac{1}{2}\|\psi'\|$. Then

$$\begin{aligned}
|(\psi - \iota(x))\varphi_\ell| &\geq |\psi'(\varphi_\ell)| - \|\psi - \iota(x) - \psi'\| \geq \frac{1}{2}\|\psi'\| - \varepsilon \\
&\geq \frac{1}{2}\|\psi - \iota(x)\| - \frac{\varepsilon}{2} - \varepsilon \geq \frac{1}{2}\|\psi - \iota(x)\| - 2\varepsilon,
\end{aligned}$$

which proves (7.12). Given $\varphi_\ell$ and $m > \ell$, since $k_n \xrightarrow{\text{weak}} x$ there exists $k_n$ with $n > m$ and such that $|\varphi_\ell(k_n - x)| < \frac{1}{m}$. Then

$$\begin{aligned}
|(\psi - \iota(x))\varphi_\ell| &\leq |(\psi - \iota(k_n))\varphi_\ell| + |\iota(k_n)\varphi_\ell - \iota(x)\varphi_\ell| \\
&= |(\psi - \iota(k_n))\varphi_\ell| + |\varphi_\ell(k_n - x)| \leq \frac{1}{n} + \frac{1}{m} < \frac{2}{n}.
\end{aligned}$$

It follows that $(\psi - \iota(x))\varphi_\ell = 0$ for all $\ell$. Combined with (7.12), we get that $\psi = \iota(x)$. $\square$

### 7.2.1. Exercises.

(7.2.1) Show that the natural embedding $\iota : \mathcal{X} \to \mathcal{X}^{**}$ given by $\iota(x) = \hat{x}$ is a linear isometry.

(7.2.2) Show that the sets $N(x_1, \ldots, x_k; \varepsilon)$, where $x_1, \ldots, x_k \in \mathcal{X}$ are linearly independent and $\varepsilon > 0$, form a base for $\sigma(\mathcal{X}^*, \mathcal{X})$.

(7.2.3) Prove Lemma 7.2.2.

(7.2.4) Prove Proposition 7.2.3.

(7.2.5) Prove Proposition 7.2.6 *(Hint: take a good look at the proof of Proposition 7.1.11)*. Show also that the completeness of $\mathcal{X}$ is crucial and cannot be dispensed with.

(7.2.6) Let $\mathcal{X}$ be a normed space and $\{\varphi_n\}$ a weak*-convergent net in $\mathcal{X}^*$ with $\varphi_n \xrightarrow{\text{weak}^*} \varphi$. Show that $\|\varphi\| \leq \liminf_n \|\varphi_n\|$. Find an example where the inequality is strict.

(7.2.7) Prove the assertions in Examples 7.2.7.

(7.2.8) Let $\mathcal{X} = \ell^1[0,1]$, so that its dual is $\mathcal{X}^* = \ell^\infty[0,1]$ (see Exercise 5.6.11), and put

$$M = \{x \in \mathcal{X}^* : \ \operatorname{supp} x \ \text{ is countable}\}.$$

Show that $M$ is closed in norm and in the weak topology, and that it is dense in the weak*-topology.

(7.2.9) Let $\mathcal{X}$ be a finite-dimensional normed space. Show that $\mathcal{X}$ is reflexive.

(7.2.10) Complete the proof of Lemma 7.2.17 by showing that if (i) is false, then $\alpha \notin \overline{\gamma(B_1^{\mathcal{X}}(0))}$.

(7.2.11) By Goldstine's Theorem the unit ball of $c_0$ is weak*-dense in the unit ball of $\ell^\infty(\mathbb{N})$. Given $f \in \ell^\infty(\mathbb{N})$ with $\|f\| \leq 1$, find a net $\{g_j\} \subset c_0$ such that $\|g_j\| = 1$ for all $j$, and $g_j \to f$ in the weak*-topology.

(7.2.12) Let $\mathcal{X}$ be a normed space, and $\varphi : \mathcal{X}^* \to \mathbb{C}$ a weak*-continuous linear map. Use the Closed Graph Theorem to show that $\varphi$ is bounded *(This is not hard, though it is not the easiest way to prove this)*.

(7.2.13) Let $\mathcal{X}$ be a normed space and $\psi \in \mathcal{X}^{**}$ nonzero. Show that there exists $\{x_j\} \subset \mathcal{X}$ with $\hat{x}_j \xrightarrow{\text{weak}^*} \psi$ and $\|x_j\| = \|\psi\|$ for all $j$.

## 7.3. Polars and Prepolars

When dealing with subsets of Hilbert spaces, a crucial concept is that of the orthogonal subspace. Not having orthogonals for arbitrary Banach spaces makes some proofs in Chapter 5 more cumbersome than their Hilbert-space counterparts. As we will see in this section, there is actually a generalization of the notion of orthogonal to Banach spaces. The leap one needs to make is that we can replace "take the inner product against a vector" with "evaluate at a linear functional".

**7.3.1. Definition.**

*Let $\mathcal{X}$ be a topological vector space, and $X \subset \mathcal{X}$. The* ***polar*** *of $X$ is the set*

$$X^o = \{\varphi \in \mathcal{X}^* : \ |\varphi(x)| \leq 1 \ \text{ for all } x \in X\}.$$

*Let $Y \subset \mathcal{X}^*$. The* ***prepolar*** *of $Y$ is the set*

$$Y_o = \{x \in \mathcal{X} : \ |\varphi(x)| \leq 1 \ \text{ for all } \varphi \in Y\}.$$

There are alternative definitions of the polar and prepolar where one uses the real part instead of the absolute value. Here we will be mostly concerned with polars and prepolars of subspaces, in which case the notions agree.

Fairly straightforward computations (Exercise 7.3.1) show that, when $X \subset \mathcal{X}$ is a subspace,

$$X^o = \{\varphi \in \mathcal{X}^* : \ \varphi(x) = 0 \ \text{ for all } x \in X\}.$$

and when $Y \subset \mathcal{X}^*$ is a subspace

$$Y_o = \{x \in \mathcal{X} : \ \varphi(x) = 0 \ \text{ for all } \varphi \in Y\}.$$

For any Banach space $\mathcal{X}$,

$$\big[B_1^{\mathcal{X}}(0)\big]^o = \overline{B_1^{\mathcal{X}^*}}(0), \qquad \big[B_1^{\mathcal{X}^*}(0)\big]_o = \overline{B_1^{\mathcal{X}}}(0).$$

(Exercise 7.3.2).

We list some elementary properties that are routine to verify.

***7.3.2. Proposition.***

*Let $\mathcal{X}$ be a topological vector space, $X \subset \mathcal{X}$, $Y \subset \mathcal{X}^*$. Then*

*(i) $X^o$ is convex, balanced, and weak*-closed;*

*(ii) $Y_o$ is convex, balanced, and weak-closed;*

*(iii) if $X_1 \subset X_2 \subset \mathcal{X}$, then $X_2^o \subset X_1^o$;*

*(iv) if $Y_1 \subset Y_2 \subset \mathcal{X}^*$, then $(Y_2)_o \subset (Y_1)_o$;*

*(v) if $Y_j \subset \mathcal{X}^*$, $j \in J$, then*

$$\Big(\bigcup_j Y_j\Big)_o = \Big(\operatorname{conv} \bigcup_j Y_j\Big)_o;$$

*(vi)* *if* $X_j \subset \mathcal{X}$, $j \in J$, *then*

$$\bigcup_j X_j^o \subset \Big(\bigcap_j X_j\Big)^o,$$

*and the reverse inclusion can fail;*

*(vii)* *if* $X_j \subset \mathcal{X}$, $j \in J$, *then*

$$\bigcap_j X_j^o = \Big(\bigcup_j X_j\Big)^o.$$

*(viii)* $(\overline{X})^o = X^o$, $(\overline{Y}^{w^*})_o = Y_o$.

*Proof.* Exercise 7.3.3. □

**7.3.3. Definition.**

*Let* $\mathcal{X}$ *be a topological vector space and* $A \subset \mathcal{X}$. *The* ***convex balanced hull*** *of* $A$ *is the smallest convex and balanced set* $\mathrm{cb}\,A$ *that contains* $A$. *Similarly, the* ***closed convex balanced hull*** *of* $A$ *is the smallest closed, convex, and balanced set* $\overline{\mathrm{cb}}A$ *that contains* $A$.

When $A$ is a subspace, the convex balanced hull of $A$ is simply $A$.

***7.3.4. Proposition.***

*Let* $\mathcal{X}$ *be a locally convex space and* $X \subset \mathcal{X}$, $Y \subset \mathcal{X}^*$. *Then*

$$(X^o)_o = \overline{\mathrm{cb}}X, \qquad (Y_o)^o = \overline{\mathrm{cb}}^{w^*}Y.$$

*Proof.* Let $x \in X$. For any $\varphi \in X^o$, we have $|\varphi(x)| \leq 1$; so $\mathcal{X} \subset (X^o)_o$. As $(X^o)_o$ is convex, balanced, and norm closed (as the norm and weak closures agree on a convex set), we get that $\overline{\mathrm{cb}}X \subset (X^o)_o$. For the converse, we will show that $\mathcal{X} \setminus \overline{\mathrm{cb}}X \subset \mathcal{X} \setminus (X^o)_o$, which then implies $(X^o)_o \subset \overline{\mathrm{cb}}X$. Fix $x_0 \in \mathcal{X} \setminus \overline{\mathrm{cb}}X$. As $\overline{\mathrm{cb}}X$ is convex and closed, we can separate via Hahn–Banach; that is, applying Theorem 5.7.18 we get $\psi \in \mathcal{X}^*$ and $d \in \mathbb{R}$ such that

$$\mathrm{Re}\,\psi(x_0) < d < \mathrm{Re}\,\psi(x), \quad x \in \overline{\mathrm{cb}}X. \tag{7.13}$$

As $\overline{\mathrm{cb}}X$ is balanced, for any $x \in X$ we have $d < \pm\mathrm{Re}\,\psi(x)$. Thus $d < 0$. Replacing $\psi$ with $\frac{1}{d}\,\psi$ the inequalities (7.13) become

$$\mathrm{Re}\,\psi(x) < 1 < \mathrm{Re}\,\psi(x_0), \quad x \in \overline{\mathrm{cb}}X. \tag{7.14}$$

For any $x \in \overline{\mathrm{cb}}X$, we have $|\psi(x)| = \psi(x)e^{i\theta} = \psi(e^{-i\theta}x)$, and $e^{-i\theta}x \in \overline{\mathrm{cb}}X$ since it is balanced. So (7.14) implies

$$|\psi(x)| < 1 < \operatorname{Re}\psi(x_0), \quad x \in \overline{\mathrm{cb}}X. \tag{7.15}$$

This means that $\psi \in X^o$. And $|\psi(x_0)| > \operatorname{Re}\psi(x_0) > 1$, so $x_0 \notin (X^o)_o$. The reverse inclusion is then established, and so $(X^o)_o = \overline{\mathrm{cb}}X$.

For $Y$ the proof runs almost the same, now taking into account that $(Y_o)^o$ is weak*-closed, and that $\mathcal{X}$ is the dual of $(\mathcal{X}^*, \sigma(\mathcal{X}^*, \mathcal{X}))$ (Proposition 7.2.10). □

***7.3.5. Corollary.***

*Let $\mathcal{X}$ be a locally convex space and $X \subset \mathcal{X}$, $Y \subset \mathcal{X}^*$ be subspaces. Then*

*(i) $X$ is dense in $\mathcal{X}$ if and only if $X^o = \{0\}$;*

*(ii) $Y$ is weak*-dense in $\mathcal{X}^*$ if and only if $Y_o = \{0\}$.*

*Proof.* Exercise 7.3.4. □

*7.3.6. Remark*. As a nice application of the ideas above, we can now make Remark 5.5.11 into an actual proof of Lemma 5.5.10. Indeed, if

$$\bigcap_{j=1}^{n} \ker\varphi_j \subset \ker\varphi,$$

we can rewrite this relation as

$$\bigcap_{j=1}^{n} \{\mathbb{C}\varphi_j\}_o \subset \{\mathbb{C}\varphi\}_o. \tag{7.16}$$

Now one notices that

$$\bigcap_{j=1}^{n} \{\mathbb{C}\varphi_j\}_o = \Big[\operatorname{span}\{\varphi_1, \ldots, \varphi_n\}\Big]_o,$$

so (7.16) becomes

$$\operatorname{span}\{\varphi_1, \ldots, \varphi_n\}_o \subset \{\mathbb{C}\varphi\}_o,$$

and then Proposition 7.3.4—with no closure due to finite-dimension—gives us

$$\{\mathbb{C}\varphi\} \subset \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}.$$

Polars behave nicely with regards to duals:

**7.3.7. Theorem.**

*Let $\mathcal{X}$ be a Banach space and $M \subset \mathcal{X}$ be a closed subspace. Then there exist canonical isometric isomorphisms*

$$(\mathcal{X}/M)^* \simeq M^o, \qquad M^* \simeq \mathcal{X}^*/M^o.$$

*Proof.* We define $\Gamma : M^o \to (\mathcal{X}/M)^*$ by $(\Gamma\varphi)(x+M) = \varphi(x)$. This is well-defined because $\varphi|_M = 0$ as $\varphi \in M^o$. Given $\varepsilon > 0$ choose $m \in M$ such that $\|x+m\| < \|x+M\| + \varepsilon$. Then, for any $\varphi \in M^o$,

$$|(\Gamma\varphi)(x+M)| = |\varphi(x)| = |\varphi(x+m)| \le \|\varphi\|\,\|x+m\| < \|\varphi\|\,(\|x+M\| + \varepsilon).$$

As this can be done for all $\varepsilon > 0$, $\|\Gamma\varphi\| \le \|\varphi\|$. At the same time, again fixing $\varepsilon > 0$, we can choose $x$ with $\|x\| = 1$ and such that $|\varphi(x)| > \|\varphi\| - \varepsilon$. Now $\|x+M\| \le \|x\| = 1$ and

$$|(\Gamma\varphi)(x+M)| = |\varphi(x)| > \|\varphi\| - \varepsilon > (\|\varphi\| - \varepsilon)\|x+M\|.$$

Thus $\|\Gamma\varphi\| \ge \|\varphi\| - \varepsilon$ and, as $\varepsilon$ was arbitrary, $\|\Gamma\varphi\| = \|\varphi\|$ and $\Gamma$ is an isometry. Checking that $\Gamma$ is linear is a straightforward affair, so it remains to check that it is surjective. Given $\psi \in (\mathcal{X}/M)^*$, we define $\tilde{\psi} : \mathcal{X} \to \mathbb{C}$ by $\tilde{\psi}(x) = \psi(x+M)$. The linearity of $\tilde{\psi}$ follows directly from that of $\psi$. We also have

$$|\tilde{\psi}(x)| = |\psi(x+M)| \le \|\psi\|\,\|x+M\| \le \|\psi\|\,\|x\|,$$

so $\tilde{\psi} \in \mathcal{X}^*$; and $\tilde{\psi}(m) = 0$ for all $m \in M$, so $\tilde{\psi} \in M^o$. Finally, $(\Gamma\tilde{\psi})(x+M) = \tilde{\psi}(x) = \psi(x+M)$ and so $\Gamma\tilde{\psi} = \psi$ and $\Gamma$ is surjective.

As for the second isomorphism, let $R : \mathcal{X}^* \to M^*$ be the restriction map $R\varphi = \varphi|_M$. Then $\ker R = M^o$. Given any $\varphi_0 \in M^*$, by Hahn–Banach (Corollary 5.7.6) there exists $\varphi \in \mathcal{X}^*$ that extends $\varphi_0$, so $R\varphi = \varphi_0$ and $R$ is surjective. By the First Isomorphism Theorem (Exercise 6.1.1) we have $\mathcal{X}^*/M^o \simeq \operatorname{ran} R = M^*$ via $\tilde{R}(\varphi + M^o) = R\varphi$. It remains to show that $\tilde{R}$ is isometric. Given $\phi_0 \in M^*$, as above we can choose $\varphi \in \mathcal{X}^*$ with $\varphi|_M = \varphi_0$ and $\|\varphi\| = \|\varphi_0\|$. For any $\psi \in M^o$,

$$\begin{aligned}\|\varphi_0\| &= \sup\{|\varphi_0(m)| :\ m \in B_1^M(0)\} = \sup\{|(\varphi_0 + \psi)(m)| :\ m \in B_1^M(0)\} \\ &\le \sup\{|(\varphi_0 + \psi)(m)| :\ m \in B_1^{\mathcal{X}}(0)\} = \|\varphi + \psi\|.\end{aligned}$$

It follows that

$$\|\varphi_0\| \le \|\varphi + M^o\| \le \|\varphi\| = \|\varphi_0\|.$$

This says that

$$\|\tilde{R}(\varphi + M^o)\| = \|R\varphi\| = \|\varphi_0\| = \|\varphi + M^o\|,$$

and $R$ is isometric. $\square$

Here is one application of the above result, that characterizes which subspaces of the dual are duals themselves.

**7.3.8. Corollary.**

*Let $\mathcal{X}$ be a normed space. Then every weak*-closed subspace of $\mathcal{X}^*$ has a predual.*

*Proof.* Fix $M \subset \mathcal{X}^*$, weak*-closed subspace. By Theorem 7.3.7, applied to $M_o \subset \mathcal{X}$,

$$(\mathcal{X}/M_o)^* \simeq (M_o)^o = \overline{M}^{w^*} = M.$$

where the equality comes from Proposition 7.3.4 and the fact that $M$ is weak*-closed. ☐

**7.3.1. Exercises.**

(7.3.1) Show that if $X \subset \mathcal{X}$ is a subspace, then

$$X^o = \{\varphi \in \mathcal{X}^* : \ \varphi(x) = 0 \ \text{ for all } x \in X\}.$$

Similarly, show that if $Y \subset \mathcal{X}^*$ is a subspace, then

$$Y_o = \{x \in \mathcal{X} : \ \varphi(x) = 0 \ \text{ for all } \varphi \in Y\}.$$

(7.3.2) Show that

$$\left[B_1^{\mathcal{X}}(0)\right]^o = \overline{B_1(0)^{\mathcal{X}^*}}, \qquad \left[B_1(0)^{\mathcal{X}^*}\right]_o = \overline{B_1(0)^{\mathcal{X}}}.$$

(7.3.3) Prove Proposition 7.3.2.

(7.3.4) Prove Corollary 7.3.5.

(7.3.5) Let $\mathcal{X}$ be a normed space, and $K \subset \mathcal{X}$ a subspace. Show that $K^{oo} = \overline{J_{\mathcal{X}}K}^{w^*}$, where $J_{\mathcal{X}}$ is the usual embedding of $\mathcal{X}$ in $\mathcal{X}^{**}$.

## 7.4. Spaces of Continuous Functions

We recall the following definitions about continuous functions. This is just a copy of Definition 2.6.1.

**7.4.1. Definition.**

*Given a locally compact Hausdorff space $T$, the* ***support*** *of $f : T \to \mathbb{C}$ is the set*

$$\operatorname{supp} f = \overline{\{t \in T : \ f(t) \neq 0\}} \subset T.$$

*The function $f$ is said to* ***vanish at infinity*** *if given $\varepsilon > 0$ there exists $K \subset T$, compact, with $|f(t)| < \varepsilon$ for all $t \in T \setminus K$.*

*We denote by $C(T)$ the set of continuous functions $f : T \to \mathbb{C}$, by $C_0(T)$ the set of continuous functions $f : T \to \mathbb{C}$ vanishing at infinity, and by $C_c(T)$ the set of continuous functions $f : T \to \mathbb{C}$ with compact support.*

We always have $C_c(T) \subset C_0(T) \subset C(T)$. In the particular case where $T$ is compact, $C_0(T)$ and $C_c(T)$ coincide with $C(T)$. By considering pointwise addition and multiplication, together with multiplication by scalars, $C_0(T)$ is a complex algebra. We also consider $C_0(T)$ as a Banach space (Exercise 7.4.1) with respect to the norm

$$\|f\| = \sup\{|f(t)| : \ t \in T\}.$$

We have that $C_c(T)$ is dense in $C_0(T)$ (Proposition 2.6.2). When $T$ is not compact, the existence of the infinity norm makes $C_0(T)$ a more suitable object for our purposes than $C(T)$, where we do not have a norm.

We first establish a kind of representation theorem for locally compact spaces (more generally, the result applies to any topological space where the result of Urysohn's Lemma holds).

**7.4.2. Proposition.**

*Let $T$ be locally compact Hausdorff. Then there exists a set $A$ such that $T$ is homeomorphic to a subspace of $[0,1]^A$, considered with the product topology.*

*Proof.* Let $A = C(T, [0,1])$ and take $h : T \to [0,1]^A$ to be given by $h(t) = (f(t))_{f\in A}$. We immediately get that $h$ is continuous, because if $t_j \to t$, then $f(t_j) \to f(t)$ for all $f \in A$, so $h(t_j) \to h(t)$. Also, $h$ is injective, for if $t \neq s$ then by Urysohn's Lemma there exists $f \in A$ such that $f(t) \neq f(s)$, so $h(t) \neq h(s)$. It remains to check that $h^{-1} : h(T) \to T$ is continuous. But if $h(t_j) \to h(t)$ this means that $f(t_j) \to f(t)$ for all $f \in A$; the proof of Corollary 2.6.7 (since it only uses functions in $A$) then shows that $t_j \to t$. Thus $h$ is injective, and bicontinuous onto $h(T) \subset [0,1]^A$. $\square$

We characterized the dual of $C_0(T)$ in Theorem 5.6.15. We will pay attention to an important subset of $C_0(T)^*$:

**7.4.3. Definition.**

*Let $T$ be locally compact Hausdorff. We say that $\varphi : C_0(T) \to \mathbb{C}$ is a* ***character*** *if $\varphi$ is nonzero, linear, and multiplicative.*

Characters exist for any $T$, as we can always consider the map $\varphi_t(f) = f(t)$ for fixed $t \in T$. Our first goal is to prove Theorem 7.4.8, and we will build up to it through a few results. Initially we will restrict to the case where $T$ is compact.

**7.4.4. Lemma.**

*Let $T$ be compact Hausdorff and $\varphi : C(T) \to \mathbb{C}$ be a character. Then, for each $f \in C(T)$, there exists $t_f \in T$ such that $\varphi(f) = f(t_f)$.*

*Proof.* We have $\varphi(1) = \varphi(1^2) = \varphi(1)^2$; this forces $\varphi(1) = 1$, as the alternative is $\varphi(1) = 0$, which would force $\varphi = 0$. So $\varphi(\lambda) = \lambda$ for all $\lambda \in \mathbb{C}$. Now if $f \in C(T)$ and $f(t) \neq 0$ for all $t$, we get by compactness that there exists $c > 0$ with $|f(t)| \geq c > 0$ for all $t$, and then $1/f \in C(T)$ (Exercise 2.6.7). So

$$1 = \varphi(f\,1/f) = \varphi(f)\,\varphi(1/f),$$

and then $\varphi(f) \neq 0$. Thus, if $\varphi(f) = 0$, there exists $y_f \in T$ with $f(y_f) = 0$.

For an arbitrary $f \in C(T)$, consider the function $f - \varphi(f)$. As $\varphi(f - \varphi(f)) = 0$, the above implies that there exists $t_f \in T$ with $f(t_f) - \varphi(f) = 0$, i.e., $f(t_f) = \varphi(f)$. $\square$

The space $C_0(T)$ carries a lot of structure. We have already mentioned that we can see it as a complex algebra. It also carries an **order**:

**7.4.5. Definition.**

*We say that $f \in C_0(T)$ is **positive**, denoted $f \geq 0$ if $f(t) \geq 0$ for all $t \in T$. This order induces itself an order in the dual: we say that $\varphi \in C_0(T)^*$ is **positive** if $\varphi(f) \geq 0$ whenever $f \geq 0$. Similarly, we say that $f$ is **real** if $f(t) \in \mathbb{R}$ for all $t \in T$, and $\varphi \in C_0(T)^*$ is real when $\varphi(f) \in \mathbb{R}$ whenever $f$ is real.*

For any real $\varphi$, since we may write $f = f_1 + if_2$ with $f_1, f_2$ real, we have that $\varphi(f) = \varphi(f_1) + i\varphi(f_2)$. In particular, $\varphi(\operatorname{Re} f) = \operatorname{Re}\varphi(f)$, $\varphi(\overline{f}) = \overline{\varphi(f)}$, and if $\varphi$ is also a character,

$$\varphi(|f|^2) = \varphi(\overline{f}f) = \overline{\varphi(f)}\varphi(f) = |\varphi(f)|^2.$$

Here are several consequences of Lemma 7.4.4.

**7.4.6. Proposition.**

*Let $T$ be compact Hausdorff and $\varphi : C(T) \to \mathbb{C}$ be nonzero, linear, and multiplicative. Then*

(i) $\varphi$ *is real;*

(ii) $\varphi$ *is positive;*

(iii) $\varphi$ *is bounded;*

(iv) *there exists a unique* $t \in T$ *such that* $\varphi(f) = f(t)$ *for all* $f \in C(T)$.

*Proof.* Fix $f \in C(T)$. By Lemma 7.4.4 there exists $t_f \in T$ with $\varphi(f) = f(t_f)$. If $f$ is real, then $\varphi(f) = f(t_f) \in \mathbb{R}$. If $f$ is positive, then $\varphi(f) = f(t_f) \geq 0$. We also get immediately that $\varphi$ is bounded:

$$|\varphi(f)| = |f(t_f)| \leq \|f\|.$$

Now, for each $f \in C(T)$, consider the set

$$T_f = \{t \in T : \ f(t) = \varphi(f)\}.$$

We know from Lemma 7.4.4 that $T_f \neq \varnothing$ for all $f$. Since $T_f = f^{-1}(\{\varphi(f)\})$, it follows that $T_f$ is closed, and so compact. Given $f_1, \ldots, f_n \in C(T)$, consider the function

$$g = \sum_{j=1}^{n} |f_j - \varphi(f_j)|^2.$$

As $\varphi$ is real, linear, and multiplicative,

$$\varphi(g) = \sum_{j=1}^{n} \varphi(f_j - \varphi(f_j))\, \varphi(\overline{f_j - \varphi(f_j)}) = 0.$$

By Lemma 7.4.4 there exists $t_0 \in T$ with $g(t_0) = \varphi(g)$; so $g(t_0) = 0$. It follows that $f_j(t_0) = \varphi(f_j)$ for all $j = 1, \ldots, n$. This shows that $\bigcap_{j=1}^n T_{f_j} \neq \varnothing$. That is, the family $\{T_f\}_{f \in C(T)}$ has the finite intersection property. By the compactness of $T$ and Proposition 1.8.19, there exists $t_1 \in \bigcap_{f \in C(T)} T_f$. That is, $\varphi(f) = f(t_1)$ for all $f \in C(T)$.

The uniqueness of $t_1$ follows from the fact that $C(T)$ separates the points in $T$ (Exercise 7.4.2). $\square$

The characters are equally described when $T$ is locally compact Hausdorff:.

***7.4.7. Proposition.***

*Let* $T$ *be locally compact Hausdorff. Then the nonzero linear, multiplicative functionals on* $C_0(T)$ *are precisely* $\{\tau_t : \ t \in T\}$, *where* $\tau_t(f) = f(t)$.

*Proof.* If $T$ is already compact, we are in the situation of Proposition 7.4.6. If $T$ is not compact, let $T_\infty$ be the one-point compactification of $T$ (Proposition 1.8.27). For $f \in C(T_\infty)$, let $\alpha = f(\infty)$. By continuity, given $\varepsilon > 0$ there exists an open neighbourhood of $\infty$, say $(T \setminus K) \cup \{\infty\}$ with $K$ compact, such that $|f(t) - \alpha| < \varepsilon$ for all $t \in T \setminus K$. We can read this as saying that the function $f - \alpha$ is less than $\varepsilon$ outside of $K$. That is, $f - \alpha \in C_0(T)$. So we can write any $f \in C(T_\infty)$ as $f = f_0 + \alpha 1$, where $f_0 \in C_0(T)$.

Given a character $\tau : C_0(T) \to \mathbb{C}$, we extend it to $C(T_\infty)$ by

$$\tilde{\tau}(f_0 + \alpha 1) = \tau(f_0) + \alpha.$$

It is straightforward to check that $\tilde{\tau}$ is linear and multiplicative; and it is nonzero since $\tau$ is nonzero. By Proposition 7.4.6, there exists $t_0 \in T_\infty$ with $\tilde{\tau}(f) = f(t_0)$ for all $f \in C(T_\infty)$. We cannot have $t_0 = \infty$ because $\tau \neq 0$ (that is, by hypothesis there exists $f \in C_0(T)$ with $\tau(f) \neq 0$ and $f(\infty) = 0$); thus $t_0 \in T$. □

The Gelfand–Kolmogorov Theorem below has far reaching consequences, both in a technical and a philosophical sense. Technically, we will use the result several times. Philosophically, the theorem is remarkable in that it relates a purely topological property (homeomorphism) with a purely algebraic one (isomorphism of algebras); this plays a big role in generalizing results about spaces of functions to non-commutative settings.

***7.4.8. Theorem. (Gelfand–Kolmogorov)***

*Let $T, S$ be locally compact Hausdorff spaces. The following statements are equivalent:*

*(i) $T$ and $S$ are homeomorphic;*

*(ii) $C_0(T)$ and $C_0(S)$ are isomorphic as algebras.*

*Proof.* Let $h : T \to S$ be a homeomorphism. Define $\Gamma : C_0(S) \to C_0(T)$ by $\Gamma f = g \circ h$. It is straightforward to check that $\Gamma$ is an algebra isomorphism.

Conversely, let $\Gamma : C_0(S) \to C_0(T)$ be an algebra isomorphism. For each $t \in T$, the map $\Gamma_t : C_0(S) \to \mathbb{C}$ given by $\Gamma_t(g) = (\Gamma g)(t)$ is a character (the fact that $\pi$ is surjective guarantees that $\Gamma_t$ is nonzero). By Proposition 7.4.7 there exists a unique $h(t) \in S$ such that

$$(\Gamma g)(t) = \Gamma_t(g) = g(h(t)), \qquad g \in C_0(S).$$

The function $\Gamma g$ is continuous, so if $t_j \to t$ then $g(h(t_j)) = (\Gamma g)(t_j) \to (\Gamma g)(t) = g(h(t))$. As this can be done for any $g \in C_0(S)$, Corollary 2.6.7 implies that $h(t_j) \to h(t)$; and this works for any $t$ and $\{t_j\}$ with $t_j \to t$ so another application of Corollary 2.6.7 gives us that $h : T \to S$ is continuous.

Now repeat the above for $\Gamma^{-1} : C_0(T) \to C_0(S)$, and find a continuous $h' : S \to T$ with

$$(\Gamma^{-1} f)(s) = f(h'(s)), \quad f \in C_0(T).$$

Then, for any $f \in C_0(T)$,

$$f(t) = (\Gamma\Gamma^{-1} f)(t) = (\Gamma^{-1} f)(h(t)) = f(h' \circ h(t)).$$

Since $C_0(T)$ separates points in $T$ (Exercise 7.4.2), we get that $h' \circ h(t) = t$. Similarly, we obtain that $h \circ h'(s) = s$ for all $s \in S$, and so $h$ is a homeomorphism. □

Theorem 7.4.8 holds with just ring isomorphism in place of algebra isomorphism, when one considers the real algebras of real-valued continuous functions; a proof appears in [Dug66, Theorem XIII.6.5], and we also refer to [GJ02, Theorem 18]. It is interesting that a somewhat weaker condition—isomorphism as rings—will imply via Theorem 7.4.8 a stronger condition, isomorphism as algebras (this can be done directly, see Exercise 7.4.3). The situation is even more noticeable when one looks at Banach–Stone (Theorem 9.4.11); it turns out that if $C(S)$ and $C(T)$ are linearly isometrically isomorphic as Banach spaces, then they are isomorphic as algebras.

The technical idea of the proof of Theorem 7.4.8 can be summarized and expanded in the following result.

***7.4.9. Proposition.***

*Let $T, S$ be locally compact Hausdorff spaces. Then*

*(i) if $\alpha : T \to S$ is continuous, then $\pi_\alpha : C_0(S) \to C_0(T)$ given by $\pi_\alpha : f \longmapsto f \circ \alpha$ is a homomorphism;*

*(ii) if $\pi : C_0(S) \to C_0(T)$ is a homomorphism with dense range, there exists $\alpha : T \to S$ continuous with $\pi = \pi_\alpha$;*

*(iii) if $\pi_\alpha$ is surjective for $\alpha : T \to S$, then $\alpha$ is injective. If $\alpha$ is injective and continuous and $T$ is compact, then $\pi_\alpha$ is surjective.*

*(iv) if $\alpha : T \to S$ continuous and surjective, then $\pi_\alpha$ is isometric. If $T$ is compact and $\pi_\alpha$ is injective, then $\alpha$ is surjective.*

*Proof.*

(i) Given $f, g \in C(S)$ and $a \in \mathbb{C}$,

$$\pi_\alpha(af + g)(t) = af(\alpha(t)) + g(\alpha(t)) = (a\pi_\alpha f + \pi_\alpha g)(t), \qquad t \in T.$$

Similarly,

$$\pi_\alpha(fg)(t) = f(\alpha(t))g(\alpha(t)) = (\pi_\alpha(f)\pi_\alpha(g))(t), \qquad t \in T.$$

So $\pi_\alpha$ is a homomorphism.

(ii) Given $t \in T$, let $\Gamma_t : C_0(S) \to \mathbb{C}$ be $\Gamma_t(g) = (\pi(g))(t)$. Being a composition of a homomorphism and a character, $\Gamma_t$ is a character on $C_0(S)$; the density of the range of $\pi$ guarantees that $\Gamma_t$ is nonzero. By Proposition 7.4.7 there exists a unique $\alpha(t) \in S$ with $(\pi(g))(t) = g(\alpha(t))$ for all $t \in T$. If $t_j \to t$ then, as $\pi(g)$ is continuous,

$$g(\alpha(t_j)) = (\pi(g))(t_j) \to (\pi(g))(t) = g(\alpha(t)), \qquad g \in C_0(S).$$

By Corollary 2.6.7 we get that $\alpha(t_j) \to \alpha(t)$. And this happens for all convergent nets in $T$, so a second use of Corollary 2.6.7 gives us that $\alpha$ is continuous.

(iii) If $\pi_\alpha$ is surjective and $\alpha(t_1) = \alpha(t_2)$, for any $f \in C(T)$ there exists $g \in C(S)$ with $f = g \circ \alpha$, and so $f(t_1) = f(t_2)$. Then $t_1 = t_2$ as continuous functions separate points (via Urysohn).

Conversely, suppose that $\alpha$ is injective, continuous, and that $T$ is compact. We take $g = f \circ \alpha^{-1}$ on $\alpha(T)$ and extend by Tietze to a continuous function on $S$; then $\pi_\alpha(g) = f$.

(iv) When $\alpha$ is surjective, Let $g \in C_0(S)$. Then

$$\begin{aligned} \|\pi_\alpha(g)\|_\infty &= \max\{|g(\alpha(t))| : \ t \in T\} \\ &= \max\{|g(s)| : \ s \in S\} = \|g\|_\infty. \end{aligned}$$

Conversely, suppose that $\alpha$ is not surjective and that $T$ is compact. Then there exists $s_0 \in S \setminus \alpha(T)$. As $\alpha(T)$ is closed, there exists $V \subset S$, open, with $s_0 \in V$ and $V \cap \alpha(T) = \varnothing$. By Urysohn's Lemma (Theorem 2.6.5) there exists $g \in C_0(S)$ with $g(s_0) = 1$ and $g|_{\alpha(T)} = 0$. Then $g \neq 0$ and $\pi_\alpha(g) = 0$, so $\pi_\alpha$ is not injective.

$\square$

*7.4.10. Remark.* The surjectivity condition in (ii) of Proposition 7.4.9 cannot be dispensed with. For instance let $S = [0,1]$, $T = [2,3] \cup \{4\}$ and $\pi : C(S) \to C(T)$ be given by $\pi g = g(0)\, 1_{[2,3]}$. No function $\alpha : T \to S$ can satisfy $\pi g = g \circ \alpha$, because we would need $g(\alpha(4)) = 0$ for all $g \in C(S)$, which is impossible (we can always use Urysohn/Tietze to construct $g \in C(S)$ with $g(\alpha(4)) = 1$).

*7.4.11. Remark .* The compactness of $T$ in the converse of (iii) in Proposition 7.4.9 cannot be dispensed with. For instance let $T = (0,1]$, $S = [0,1]$ and $\alpha(t) = t$. Then $\alpha$ is continuous and injective. But $\pi_\alpha$ is not surjective. Indeed, consider $f \in C_0(T)$ given by $f(t) = \sin \frac{1}{t}$. If $\pi_\alpha g = f$ for $g \in C(S) = C[0,1]$, we need $\sin \frac{1}{t} = f(t) = g(\alpha(t)) = g(t)$ for all $t \in (0,1]$; but no such function can exist, as $f$ is not of bounded variation.

*7.4.12. Remark*. The compactness of $T$ in the converse of (iv) in Proposition 7.4.9 cannot be dispensed with. Let $T = (0,1)$, $S = [0,1]$ and $\alpha : T \to S$ the identity map. Then $\alpha$ is not surjective. But if $\pi_\alpha(g) = 0$ for some $g \in C[0,1]$, then $g(t) = 0$ for all $t \in (0,1)$; as $g$ is continuous $g = 0$. So $\pi_\alpha$ is injective, even though $\alpha$ is not surjective.

Besides that argument provided earlier in this section, there are other well-known ways to prove that characters on $C(T)$ are point evaluations. One way is, via Riesz–Markov (Theorem 5.6.15), to look at $\varphi$ as a measure and show that a measure that gives rise to a multiplicative integral has to be a Dirac measure (Exercise 2.10.19; one would first need something like Lemma 7.4.4 to show that a character is bounded in order to apply the theorem). And yet another way (Exercise 9.2.11), from the theory of Banach algebras, is to notice that $\ker\varphi$ is a maximal ideal precisely when $\varphi$ is a character, and use the following:

**7.4.13. Proposition.**

*Let $T$ be locally compact Hausdorff, and $J \subset C_0(T)$ a closed ideal. Then there exists $T_0 \subset T$, closed, such that*

$$J = \{f \in C_0(T) : \ f|_{T_0} = 0\}. \tag{7.17}$$

*Proof.* We assume first that $T$ is compact. Let $T_0 = \{t \in T : \ f(t) = 0 \ \text{ for all } f \in J\}$. Note that

$$T_0 = \bigcap_{f \in J} f^{-1}(\{0\}),$$

so $T_0$ is closed and thus compact. Write $J_0 = \{f \in C_0(T) : \ f|_{T_0} = 0\}$. It is clear that $J \subset J_0$.

Let $V \subset T$ be open with $T_0 \subset V$. For any $g \in J$, we have $|g|^2 = \bar{g}\, g \in J$. Given $t \in T \setminus V$, there exists $g_t \in J$ with $g_t(t) \neq 0$; by replacing $g_t$ with $|g_t|^2 \in J$, we may assume that $g_t(t) > 0$. Then there exists an open neighbourhood $V_t$ of $t$ such that $g_t > 0$ in all of $V_t$. As $T \setminus V \subset \bigcup_t V_t$ and it is compact, we get $t_1, \ldots, t_s$ such that $V_{t_1}, \ldots, V_{t_s}$ covers all of $T \setminus V$. Put $g = g_{t_1} + \cdots + g_{t_s} \in J$. As $g > 0$ in all of the compact set $T \setminus V$, $g_2 = 1/g \in C(T \setminus V)$ (Exercise 2.6.7). Using again that $T \setminus V$ is compact, we get from Tietze's Extension Theorem 2.6.9 an extension, that we still call $g_2$, to all of $T$. Let $g_V = g\, g_2 \in J$ (since $g \in J$). We have $g_V|_{T\setminus V} = 1$, $g_V|_{T_0} = 0$. We can prescribe $\|g_V\| \leq 2$ in the following way. For each $t \in T\setminus V$ there exists by continuity an open set $U_t$ such that $|g_V| \leq 2$ on $U_t$. Taking $U = \bigcup_{t\in T\setminus V} U_t$ we get an open set such that $T \setminus V \subset U$ and $|g_V| \leq 2$ on $U$.

Now Urysohn's Lemma produces $h \in C_0(T)$ such that $\|h\| \leq 1$, $h|_{T\setminus V} = 1$ and $\operatorname{supp} h \subset U$. Then $hg_V \in J$ and $\|hg_V\| \leq 2$, so we can replace $g_V$ with $hg_V$.

Now let $f \in J_0$, and fix $\varepsilon > 0$. For each $t \in T_0$ we have $f(t) = 0$, so there exists an open set $W_t$ with $t \in W_t$ and $|f| < \varepsilon$ on $W_t$. As $T_0$ is compact, it is covered by finitely many $W_{t_1}, \ldots, W_{t_n}$. Let $W = \bigcup_{j=1}^n W_{t_j}$, open. We have $T_0 \subset W$ and $|f| < \varepsilon$ on $W$. Let $g_W$ as in the previous paragraph. Using again that $J$ is an ideal, $fg_W \in J$.

For any $t \in T \setminus W$, since $g_W(t) = 1$ we get $f(t) - f(t)g_W(t) = 0$. For $t \in V$ we have $|f(t) - f(t)g_W(t)| = |f(t)|\,|1 - g_W(t)| \leq 2\varepsilon$. Thus

$$\|f - fg_W\| \leq 2\varepsilon.$$

As we can do this for any $\varepsilon > 0$, we have shown that $f \in \overline{J} = J$. So $J_0 \subset J$, and the equality follows.

For the general case, let $T_\infty$ be the one-point compactification of $T$ (Proposition 1.8.27). Since $C_0(T)$ is an ideal in $C(T_\infty)$, $J$ is also an ideal in $C(T_\infty)$. By the previous part of the proof there exists $\tilde{T}_0 \subset T_\infty$, closed, such that $J = \{f \in C(T_\infty) :\ f|_{\tilde{T}_0} = 0\}$. Since $f(\infty) = 0$ for all $f \in J$ necessarily $\infty \in \tilde{T}_0$. The set $T_0 = \tilde{T}_0 \setminus \{\infty\}$ is closed in $T$ (because $T \setminus T_0 = T_\infty \setminus \tilde{T}_0$ is open) and so $J$ is of the form (7.17). $\square$

We have already seen a couple of examples of how useful the one-point compactification of a locally compact Hausdorff space can be. The one-point compactification can be performed on a non-locally compact space, but at the risk of the new topology not being Hausdorff; precisely when $T$ is Hausdorff and locally compact (also called **Tychonoff**) the one-point compactification is Hausdorff. So, when talking about compactifications, it makes sense to restrict ourselves to locally compact Hausdorff spaces.

One-point compactifications usually lack a very basic property, which is the possibility of extending continuous bounded functions. The canonical example of this situation is the space $(0, 1]$, with the obvious compactification $[0, 1]$, and the function $f(t) = \sin \frac{1}{t}$. Although continuous and bounded, this function cannot be extended to a continuous function on the closure $[0, 1]$, as it is not of bounded variation. This problem can be solved via a larger compactification, as we will see right now.

Recall that we denote by $C_b(T)$ the set of all complex-valued continuous bounded functions on $T$. It carries a canonical structure as a Banach algebra with the norm $\|f\| = \sup\{|f(t)| :\ t \in T\}$.

**7.4.14. Theorem.** *(Stone–Čech Compactification)*

*Let $T$ be a locally compact Hausdorff space. Then there exists a compact Hausdorff space $\beta T$ such that*

*(i) there exists $\delta : T \to \beta T$ continuous, such that it is a homeomorphism onto its image;*

*(ii) $\delta(T)$ is dense in $\beta T$;*

*(iii) for any $f \in C_b(T)$, there exists a unique $\tilde{f} \in C(\beta T)$ with $\tilde{f} \circ \delta = f$.*

*Any compact space satisfying the above three properties is homeomorphic to $\beta T$. More generally, $\beta T$ satisfies the following universal property: if $S$ is a compact Hausdorff space and there exists $f : T \to S$ continuous,*

$$\begin{tikzcd} T \arrow[r, "\delta"] \arrow[dr, "f"'] & \beta T \arrow[d, dashed, "\beta f"] \\ & S \end{tikzcd} \tag{7.18}$$

*then there exists $\beta f : \beta T \to S$, continuous, with $\beta f \circ \delta = f$.*

*Proof.* Let $\delta : T \to C_b(T)^*$ given by $t \mapsto \delta_t$, where $\delta_t$ is the point evaluation $f \mapsto f(t)$, $f \in C_b(T)$. Since $\|\delta_t\| = 1$, the set $\delta(T)$ is a bounded subset of $C_b(T)^*$.

Let $\beta T$ be the weak*-closure of $\delta(T)$ in $C_b(T)^*$. Since it is a bounded and weak*-closed set, it is compact by Banach–Alaoglu (Theorem 7.2.13).

First, we need to prove that $\delta$ is continuous. Since the topology we are using in $C_b(T)^*$ is the weak*-topology, if $t_j \to t$ in $T$, then for any $f \in C_b(T)$ we have $\delta_{t_j}(f) = f(t_j) \to f(t) = \delta_t(f)$, so $\delta_{t_j} \xrightarrow{\text{weak}^*} \delta_t$ in $C_b(T)^*$.

Conversely, if $\delta_{t_j} \to \delta_t$, this means that $f(t_j) \to f(t)$ for all $f \in C_b(T)$. Then Corollary 2.6.7 implies that $t_j \to t$, and so $\delta$ is bicontinuous. If $\delta_{t_1} = \delta_{t_2}$, this means that $f(t_1) = f(t_2)$ for all $f \in C_b(T)$. But as $T$ is locally compact and Hausdorff, this means that $t_1 = t_2$ (Urysohn's Lemma guarantees that $C_c(T)$ separates points), so $\delta$ is injective.

The density of $\delta(T)$ in $\beta T$ occurs by definition.

If $f \in C_b(T)$, define $\tilde{f} : \beta T \to \mathbb{C}$ by $\tilde{f}(\varphi) = \varphi(f)$ (using that $\beta T \subset C_b(T)^*$). Then, $\tilde{f}$ is continuous: if $\varphi_j \xrightarrow{\text{weak}^*} \varphi$ in $\beta T$, then $\tilde{f}(\varphi_j) = \varphi_j(f) \to \varphi(f) = \tilde{f}(\varphi)$ (using once again Proposition 7.2.3). Also, for any $t \in T$, $\tilde{f}(\delta_t) = \delta_t(f) = f(t)$. The uniqueness is immediate since two continuous functions that agree on a dense subset are equal.

Now consider $S$ and $f : T \to S$ as in the diagram. By Proposition 7.4.2 there exists a homeomorphic embedding $h : S \to [0,1]^A$ for some set $A$. For

each $\gamma \in A$, let
$$f_\gamma : T \xrightarrow{f} S \xrightarrow{h} [0,1]^A \xrightarrow{\pi_\gamma} [0,1],$$
where $\pi_\gamma((b_a)) = b_\gamma$ is the projection onto the corresponding component. All three functions are continuous, so $f_\gamma$ is continuous. By the previous paragraph there exists $\tilde{f}_\gamma : \beta T \to [0,1]$, continuous, that extends $f_\gamma$. Let $g : \beta T \to [0,1]^A$ be given by $g(x) = (\tilde{f}_\gamma(x))_{\gamma \in A}$. We have (see Proposition 1.8.10)
$$g(\beta T) = g(\overline{\delta(T)}) \subset \overline{g(\delta(T))} = \overline{(f_\gamma(T))_{\gamma \in A}} = \overline{h \circ f(T)} \subset \overline{h(S)} = h(S).$$
So $\beta f = h^{-1} \circ g : \beta T \to S$ is continuous and
$$\beta f(\delta(t)) = h^{-1}(g(\delta(t))) = h^{-1}((f_\gamma(t))_{\gamma \in A}) = h^{-1}(h(f(t))) = f(t).$$

It only remains to prove the uniqueness of $\beta T$. So assume that $K$ is another compact set satisfying

(i) there exists $\gamma : T \to K$ continuous, such that it is a homeomorphism onto its image;

(ii) $\gamma(T)$ is dense in $K$;

(iii) for any $g \in C_b(T)$, there exists a unique $g' \in C(K)$ with $g' \circ \gamma = g$.

By the previous paragraph, $K$ and $\gamma$ satisfy the universal property. Applying the diagram twice, with $\beta T$ and $K$ taking the place of $S$ in each case, there exist continuous functions $\xi : \beta T \to K$ and $\eta : K \to \beta T$ that satisfy $\xi \circ \delta = \gamma$, $\eta \circ \gamma = \delta$. Then
$$\xi \circ \eta \circ \gamma = \xi \circ \delta = \gamma, \qquad \eta \circ \xi \circ \delta = \eta \circ \gamma = \delta.$$
So the continuous function $\xi \circ \eta$ is the identity on $\gamma(T)$, and $\eta \circ \xi$ is the identity on $\delta(T)$. As these are dense in $K$ and $\beta T$ respectively, $\xi$ and $\eta$ are inverses of each other and hence $K$ and $\beta T$ are homeomorphic. □

*7.4.15. Example.* It is well-known what $\beta\mathbb{N}$ is; this is one of the very few concrete examples where the Stone-Čech compactification can be calculated. We refer to Example 11.2.9 for details; see also Corollary 7.4.35.

We will need the following well-known Real Analysis result:

***7.4.16. Proposition. (Abel)***

*Let* $f(t) = \sum_{n=0}^{\infty} a_n t^n$ *with* $a_n \in \mathbb{R}$. *If* $\sum_{n=0}^{\infty} a_n$ *converges, then* $f$ *is continuous from the left at* $1$, *namely*

$$\lim_{t \to 1^-} f(t) = \sum_{n=0}^{\infty} a_n.$$

*Proof.* By changing $a_0$ if needed, we may assume without loss of generality that $f(1) = \sum_{n=0}^{\infty} a_n = 0$. If $s_n = \displaystyle\sum_{k=0}^{n} a_k$, then $a_n = s_n - s_{n-1}$, and

$$f(t) = a_0 + \sum_{n=1}^{\infty} (s_n - s_{n-1}) t^n = \sum_{n=0}^{\infty} s_n t^n - \sum_{n=0}^{\infty} s_n t^{n+1} = (1-t) \sum_{n=0}^{\infty} s_n t^n.$$

Fix $\varepsilon > 0$. Choose $n_0$ such that $|s_n| < \varepsilon$ whenever $n \geq n_0$ (recall that $s_n \to 0$ by hypothesis). Then, for any $t < 1$,

$$\left| (1-t) \sum_{n=n_0}^{\infty} s_n t^n \right| \leq \varepsilon (1-t) \sum_{n=n_0}^{\infty} t^n = \frac{\varepsilon (1-t)\, t^{n_0}}{1-t} = \varepsilon\, t^{n_0} \leq \varepsilon.$$

So $\lim_{t \to 1^-} f(t) = 0 = f(1)$. $\square$

*7.4.17. Remark* . Let us venture briefly into calculus. We need the Taylor expansion for the square root. Using the binomial expansion of $(1-t)^{1/2}$ around 0,

$$(1-t)^{1/2} = \sum_{n=0}^{\infty} \binom{1/2}{n} (-1)^n t^n, \qquad |t| < 1.$$

Working on the coefficients, when $n \geq 1$,

$$\begin{aligned}(-1)^n \binom{1/2}{n} &= \frac{(-1)^n}{n!} \prod_{k=0}^{n-1} \left(\tfrac{1}{2} - k\right) = \frac{(-1)^n}{n!} \frac{(-1)^{n-1}}{2^n} \prod_{k=1}^{n-1} (2k-1) \\ &= -\frac{1}{n!} \frac{1}{2^n} \frac{(2n-2)!}{2^{n-1}(n-1)!} = -\frac{2(2n-2)!}{4^n n!(n-1)!}.\end{aligned}$$

Thus

$$(1-t)^{1/2} = 1 - \sum_{n=1}^{\infty} \frac{2(2n-2)!}{4^n n!(n-1)!} t^n, \qquad |t| < 1 \tag{7.19}$$

The ratio test shows that the series above converges for $|t| < 1$; thus it converges uniformly on $(-r, r)$ for any $r \in (0,1)$, but we will get more than that.

Using Stirling's inequality
$$\sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n} \le n! \le \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n+1},$$
we have
$$\begin{aligned}
|a_n| = \frac{2(2n-2)!}{4^n n!(n-1)!} &\le \frac{2e\sqrt{2\pi}\,(2n-2)^{2n-2+\frac{1}{2}}\, e^{-2n+3}}{4^n\,\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}\,\sqrt{2\pi}(n-1)^{n-1+\frac{1}{2}} e^{-n+1}} \\
&\le \frac{2e\sqrt{2\pi}\,(2n-2)^{2n-2+\frac{1}{2}}\, e^{-2n+3}}{4^n\, 2\pi(n-1)^{n+\frac{1}{2}}\,(n-1)^{n-1+\frac{1}{2}} e^{-2n+1}} \\
&= \frac{e^2\, 2^{2n-\frac{1}{2}}(n-1)^{2n-\frac{3}{2}}}{4^n\,\sqrt{2\pi}(n-1)^{2n}} = \frac{e^2}{2\sqrt{\pi}}\,\frac{1}{(n-1)^{3/2}}.
\end{aligned}$$
So $\sum_n |a_n|$ converges by comparison. It follows that $\sum_n a_n t^n$ converges uniformly on $[-1,1]$. By Abel's Theorem (Proposition 7.4.16), $\sum_n a_n = (1-1)^{1/2}$=0. Thus $1 - \sum_n a_n t^n = (1-t)^{1/2}$ for all $t \in [-1,1]$, with uniform convergence. Since $\sum_n a_n$ converges absolutely, we can also apply Abel's Theorem at $t=-1$, giving us
$$\sqrt{2} = (1+1)^{1/2} = 1 - \sum_{n=1}^{\infty} \frac{(-1)^n 2(2n-2)!}{4^n n!(n-1)!}.$$

We record a consequence of the above manipulations, that we will need further down the road.

***7.4.18. Lemma.***

*Let $m \in \mathbb{N}$, $m \ge 2$. Then*
$$\sum_{k=0}^{m} \binom{1/2}{k}\binom{1/2}{m-k} = 0. \tag{7.20}$$

*Proof.* For $t \in [-1,1]$, since
$$(1+t)^{1/2} = \sum_{k=0}^{\infty} \binom{1/2}{k} t^k$$
with uniform convergence, we can exchange series below. We have
$$1 + t = [(1+t)^{1/2}]^2 = \sum_{k,j=0}^{\infty} \binom{1/2}{k}\binom{1/2}{j} t^{k+j} = \sum_{m=0}^{\infty}\left[\sum_{k=0}^{m}\binom{1/2}{k}\binom{1/2}{m-k}\right] t^m.$$
Comparing coefficients, we get (7.20). □

It is interesting to notice that using $1-t$ instead in the proof of Lemma 7.4.18 we get the identities

$$\sum_{k=0}^{m}(-1)^k\binom{1/2}{k}\binom{1/2}{m-k}=0, \qquad m\geq 2. \tag{7.21}$$

The equalities (7.20) and (7.21) can be seen as combinatorial relations, expressed in terms of the **Catalan numbers**.

Our next goal is to state and prove the Stone–Weierstrass Theorem. See Remark 7.4.21 for some details on the history of this result. We say that $A\subset C(T)$ is **selfadjoint** if for any $f\in A$ we have $f^*\in A$, where $f^*(t)=\overline{f(t)}$. We will put part of the argument in the following lemma. Given real functions $f,g$ we consider the functions $f\vee g$ and $f\wedge g$ given by

$$f\vee g(t)=\max\{f(t),g(t)\}, \quad f\wedge g(t)=\min\{f(t),g(t)\}.$$

A consequence of Lemma 7.4.19 is that both are continuous when $f,g$ are; of course, this can be easily proven on its own. These functions have already been considered on page 120 and other parts of Chapter 2, mostly in the form $f^+=f\vee 0$, $f^-=-(f\wedge 0)$.

***7.4.19. Lemma.***

*Let $T$ be compact Hausdorff and $A\subset C(T)$ a closed, unital selfadjoint subalgebra. Then, for $f,g\in A$, the functions*

$$|f|,\ |f|^{1/2},\ f\vee g,\ f\wedge g,$$

*are in $A$.*

*Proof.* Assume first that $0\leq f\leq 1$. Let $h=1-f\in A$, so $0\leq h\leq 1$. Using the Taylor series as in Remark 7.4.17, we have

$$f(t)^{1/2}=(1-h(t))^{1/2}=\sum_{n=0}^{\infty}a_nh^n(t).$$

Now

$$\begin{aligned}\Big\|f^{1/2}-\sum_{n=0}^{N}a_nh^n\Big\| &= \sup_{t\in T}\Big|f(t)^{1/2}-\sum_{n=0}^{N}a_nh^n(t)\Big|\\ &\leq \sup_{r\in[0,1]}\Big|r^{1/2}-\sum_{n=0}^{N}a_n(1-r)^n\Big|\end{aligned}$$

By the uniform convergence shown in Remark 7.4.17, given $\varepsilon>0$ there exists $N$ such that $\left\|f^{1/2}-\sum_{n=1}^{N}a_nh^n\right\|<\varepsilon$. As $\sum_{n=0}^{N}a_nh^n\in A$ and $A$ is closed, we get that $f^{1/2} \in A$. For $f \in A$ with $f \geq 0$, the above shows that $f^{1/2} = \|f\|^{1/2} \left(f/\|f\|\right)^{1/2} \in A$.

Now for arbitrary $f \in A$, using that $A$ is a selfadjoint algebra,

$$|f| = (f\, f^*)^{1/2} \in A.$$

And

$$f \vee g = \tfrac{1}{2}\,(f + g + |f - g|) \in A, \quad f \wedge g = \tfrac{1}{2}\,(f + g - |f - g|) \in A. \qquad \square$$

**7.4.20. Theorem. (Stone–Weierstrass)**

*Let $T$ be a compact Hausdorff space, and $A \subset C(T)$ a unital selfadjoint subalgebra that separates the points of $T$. Then $A$ is dense in $C(T)$.*

*Proof.* Fix $\varepsilon > 0$ and $f \in C(T)$. Assume for now that $f$ is real. Let $s, t \in T$. As $A$ separates points, there exists $h \in A$ with $h(s) \neq h(t)$. Define

$$f_{s,t}(w) = f(t) + (f(s) - f(t))\,\frac{h(w) - h(t)}{h(s) - h(t)}.$$

Since $A$ is a subalgebra (in particular, a subspace) that contains the constant functions, $f_{s,t} \in A$. Besides, we have that $f_{s,t}(s) = f(s)$, $f_{s,t}(t) = f(t)$. We may assume that $f_{s,t}$ is real, because if it isn't we replace it with $\frac{1}{2}\,(f_{s,t} + f_{s,t}^*)$, still in $A$. Define

$$U_{s,t} = \{w \in T :\ f_{s,t}(w) - f(w) < \varepsilon\}.$$

These sets are open since $f_{s,t} - f$ is continuous; we also have $s, t \in U_{s,t}$. In particular, for each $s$ we have $T \subset \bigcup_{t\in T} U_{s,t}$; by compactness, there exist $t_{s,1}, \ldots, t_{s,n(s)} \in T$ with $T \subset \bigcup_{j=1}^{n(s)} U_{s,t_{s,j}}$. Now define $h_s \in C(T)$ by

$$h_s = \bigwedge_{j=1}^{n(s)} f_{s,t_{s,j}}.$$

By Lemma 7.4.19, $h_s \in A$. Also, $h_s(s) = f(s)$, and $h_s < f + \varepsilon$. Define

$$V_s = \{w \in T :\ h_s(w) > f(w) - \varepsilon\}.$$

Because $h_s$ and $f$ are continuous, $V_s$ is open. Since $s \in V_s$, we get the inclusion $T \subset \bigcup_{s\in T} V_s$; so there exist $s_1, \ldots, s_m$ with $T \subset \bigcup_{j=1}^{m} V_{s_j}$. Let

$$g = \bigvee_{j=1}^{m} h_{s_j}.$$

Then $g \in A$ by Lemma 7.4.19, and $f - \varepsilon < g < f + \varepsilon$, so $\|f - g\| < \varepsilon$.

For general $f$ write $f = f_1 + i f_2$, where $f_1 = \frac{1}{2}\,(f + f^*)$ and $f_2 = \frac{1}{2i}\,(f - f^*)$. Then $f_1, f_2$ are real and by the above we can find, given $\varepsilon > 0$,

functions $g_1, g_2 \in A$ with $\|f_j - g_j\| < \varepsilon/2$; this gives
$$\|f - (g_1 + ig_2)\| \leq \|f_1 - g_1\| + \|f_2 - g_2\| < \varepsilon.$$
As $\varepsilon$ was arbitrary, we get that $A$ is dense. $\square$

*7.4.21. Remark*. The original version of Theorem 7.4.20 is due to Weierstrass in 1885, in the case where the algebra is $C[a, b]$, real valued, and $A$ is the algebra of polynomials. The original proof used convolution techniques similar to those in Proposition 2.8.22. In 1911, Bernstein gave a proof with an explicit algorithm, also in the spirit of convolutions, to construct the sequence of polynomials approximating $f$. The general result is from Stone [**Sto37**], who improved his proof in [**Sto48**].

*7.4.22. Remark*. The "selfadjoint" requirement in Theorem 7.4.20 is tautological when the functions are real valued, but it is essential when the functions are complex valued. Let us see a couple examples of how this can happen. Consider first the algebra $C(\overline{\mathbb{D}})$. We will see that the algebra of polynomials, which is unital and it separates points, is not dense. Take $g(z) = \bar{z}$. The equality
$$\langle z^n, \bar{z}\rangle = \int_{\mathbb{D}} z^n z\, dz = \int_{\mathbb{D}} z^{n+1}\, dz = 0$$
shows that $g$ is orthogonal in $L^2(\mathbb{D})$ to the subspace of polynomials. This implies in particular that it cannot be a uniform limit of polynomials.

For a second example consider the algebra $C(K)$, where $K = \{z \in \mathbb{C} : 1 \leq |z| \leq 2\}$. It turns out that also on this algebra the algebra of polynomials, while unital and separating points, is not dense. Let $g(z) = \frac{1}{z}$. If $\gamma$ is the circle of radius $\frac{3}{2}$ around the origin, we have $\int_\gamma \frac{1}{z}\, dz = 2\pi i$. Meanwhile, for any polynomial $p$ we have $\int_\gamma p(z)\, dz = 0$, and this implies that $g$ cannot be a uniform limit of polynomials. This looks contradictory because one is used to see holomorphic functions (which $g$ is) as uniform limits of polynomials; but this need not happen globally when the region is not simply connected (or the equivalent condition $\mathrm{Ind}_\gamma(z) = 0$ for all $z \in \mathbb{C} \setminus V$ that was used in Section 3.9).

Although Theorem 7.4.20 uses compactness in an essential way, the result can be extended to locally compact spaces with just a minimal change.

***7.4.23. Corollary.***

*Let $T$ be a locally compact, non-compact, Hausdorff space, and $\mathcal{A} \subset C_0(T)$ a selfadjoint subalgebra that separates the points of $T$ and vanishes nowhere. Then $\mathcal{A}$ is dense in $C_0(T)$.*

*Proof.* Let $T_\infty$ be the one-point compactification of $T$. We can identify $C_0(T)$ with $\{f \in C(T_\infty) : \ f(\infty) = 0\}$. Consider the algebra $\mathcal{A} + \mathbb{C}\,1 \subset C(T_\infty)$. It clearly separates the points of $T_\infty$ (because it vanishes nowhere, which we need to separate any point from $T$ and $\infty$), it contains adjoints, and it's unital; by Theorem 7.4.20, we have that $\mathcal{A} + \mathbb{C}\,1$ is dense in $C(T_\infty)$. Now take $f \in C_0(T)$ and $\varepsilon > 0$; then there exists $g \in \mathcal{A}$ and $c \in \mathbb{C}$ such that $\|f - (g + c\,1)\| < \varepsilon$. In particular

$$\varepsilon > |f(\infty) - g(\infty) - c| = |c|.$$

Then, for any $t \in T$,

$$|f(t) - g(t)| \leq |f(t) - g(t) - c| + |c| \leq \|f - (g + c\,1)\| + |c| \leq 2\varepsilon.$$

Thus $\|f - g\| < 2\varepsilon$, and as $\varepsilon$ was arbitrary we conclude that $\mathcal{A}$ is dense in $C_0(T)$. ☐

Next, an application to metrizability.

***7.4.24. Proposition.***

*Let $T$ be compact Hausdorff. Then $C(T)$ is separable if and only if $T$ is metrizable.*

*Proof.* Assume $C(T)$ is separable; so its unit ball is separable and it has a dense countable subset $\{f_n\}$. Define, on $T$, the distance

$$d(t,s) = \sum_{n=1}^\infty 2^{-n}\,|f_n(t) - f_n(s)|. \tag{7.22}$$

Note that $t_j \to t$ if and only if $f(t_j) \to f(t)$ for all $f \in C_0(T)$ (Corollary 2.6.7). One can check that $d$ is a metric and that it induces the topology of $T$ (Exercise 7.4.4).

Conversely, suppose that $T$ is metrizable, with metric $d$. This implies that $T$, being compact, is separable (Exercise 7.4.5); say $\{t_n\}$ is a dense subset. For each $n$, consider the function $d_n(t) = d(t, t_n)$. The reverse triangle inequality shows that $d_n \in C(T)$:

$$|d_n(t) - d_n(s)| = |d(t,t_n) - d(s,t_n)| \leq d(s,t).$$

Let $A$ be the algebra generated by 1, and $\{d_n\}_n$. Then $A$ is separable. It also separates points: if $d_n(t) = d_n(s)$ for all $n$, we have

$$d(t,s) \leq d(t,t_n) + d(t_n,s) = d_n(t) + d_n(s) = 2d_n(t) = 2d(t_n,t)$$

for all $n$, so $s = t$. It contains adjoints by construction, since the $d_n$ are real-valued. By Stone–Weierstrass (Theorem 7.4.20), $A$ is dense in $C(T)$ and so $C(T)$ is separable. □

Proposition 7.4.24 cannot be generalized to locally compact spaces. The implication "$C_0(T)$ separable $\implies$ $T$ metrizable" in Proposition 7.4.24 holds when $T$ is locally compact, with the same proof; but the converse does not. The problem is that there are locally compact metrizable spaces which are not separable. For example let $T = \mathbb{R}$ with the discrete topology. It is metrizable with the distance $d(s,t) = \delta_{s,t}$, and it is locally compact since each point is its own clopen compact neighbourhood. But $C_0(T)$ is not separable, for in the uncountable family $\{1_{\{t\}}\}_{t\in\mathbb{R}}$ the distance between any two elements is $\|1_{\{t\}} - 1_{\{s\}}\|_\infty = 1$.

We mention here that in the easiest example of an infinite discrete set, $\mathbb{N}$, we have $C_0(\mathbb{N}) = c_0$ (Exercise 7.4.8).

**7.4.25. Definition.**

*Let $T$ be a compact Hausdorff topological space. We denote by $C_\mathbb{R}(T)$ the real subspace of real-valued continuous functions on $T$. A subset $\mathcal{R} \subset C_\mathbb{R}(T)$ is* ***bounded*** *if there exists $g \in C_\mathbb{R}(T)$ such that $f \leq g$ for all $f \in \mathcal{R}$, meaning that $f(t) \leq g(t)$ for all $t \in T$ and $f \in \mathcal{R}$.*

*The space $C(T)$ is* ***order-complete*** *if every bounded subset $\mathcal{R} \subset C_\mathbb{R}(T)$ admits a least upper bound in $C_\mathbb{R}(T)$.*

An immediate consequence of the definition is that when the space $C(T)$ is order-complete any bounded below real-valued subset of $C(T)$ admits an infimum. The supremum/infimum will often not agree with the pointwise supremum/infimum of the functions in $\mathcal{R}$, as it has to be continuous.

It is not very common for $C(T)$ to be order-complete. We can see that $C[0,1]$ is not order-complete by considering the set $\mathcal{R} = \{f \in C([0,1] : \ 0 \leq f \leq 1, \ f|_{\left[\frac{1}{2},1\right]} = 0\}$. Then $\mathcal{R}$ is bounded but it does not have a supremum in $C[0,1]$ (Exercise 7.4.9). This should be seen as a strong suggestion that $T$ has to be very particular for $C(T)$ to be order-complete. Indeed, consider this definition:

**7.4.26. Definition.**

*A topological space $T$ is* ***extremally disconnected*** *if the closure of every open set is open (hence clopen). A compact Hausdorff extremally disconnected space is said to be* ***Stonean****.*

The property of having closures of open sets be open might look surprising, but it is not entirely unusual. For instance any discrete space is extremally disconnected, since every subset is open. So finite discrete spaces are Stonean. We will see in Corollary 7.4.35 that $\beta T$ is extremally disconnected for any discrete space $T$. For a different point of view, in Example 11.2.9 we wil show that $C(\beta\mathbb{N}) \simeq \ell^\infty(\mathbb{N})$ isometrically (in fact, not only as Banach spaces but as C*-algebras). The pointwise order is the same in both, then. In $\ell^\infty(\mathbb{N})$ the supremum of a bounded set agrees with the pointwise supremum, and hence $C(\beta\mathbb{N})$ is order-complete. Via Proposition 7.4.27 below we get another proof that $\beta\mathbb{N}$ is extremally disconnected. This gives us an example of an extremally disconnected compact Hausdorff space that is neither finite nor discrete (it is not discrete because, being a completion of $\mathbb{N}$, it has convergent non-constant nets).

We have

***7.4.27. Proposition.***

*Let $T$ be a compact Hausdorff space. The following statements are equivalent:*

*(i) $C(T)$ is order-complete;*

*(ii) $T$ is extremally disconnected.*

*Proof.* Suppose first that $C(T)$ is order-complete. Let $V \subset T$ be open and put $\mathcal{R} = \{f \in C(T) : \ f \geq 0, \ f|_V \geq 1\}$. This set is bounded below by 0, so it admits an infimum $h$. Using Urysohn's Lemma we can construct $f \in C(T)$ with $0 \leq f \leq 1$ and $f|_{\overline{V}} = 1$, so $h|_V \leq 1$. If $h(v) < 1$ for some $v \in V$, another use of Urysohn's Lemma would allow us to construct a lower bound for $\mathcal{R}$ greater than $h$, a contradiction. So $h|_{\overline{V}} = 1$. Similarly, $h|_{T\setminus\overline{V}} = 0$, for if $h(t) < 0$ for some $t \in T \setminus \overline{V}$, we can use Urysohn's Lemma to construct a lower bound greater than $h$; and we know that $h|_{T\setminus\overline{V}} \leq 0$ for given any $w \in T \setminus \overline{V}$ we can use Urysohn to construct $f \in \mathcal{R}$ with $f(v) = 0$. So $h = 1_{\overline{V}}$, and continuous. Then $\overline{V} = h^{-1}(\{1\}) = h^{-1}(1 - 1/2, 1 + 1/2)$ is open.

Conversely, suppose that $T$ is extremally disconnected. Let $\mathcal{R} \subset C_{\mathbb{R}}(T)$ be a bounded subset, and let $h$ be the function

$$h(t) = \sup\{f(t) : \ f \in \mathcal{R}\}.$$

By Corollary 1.8.13, $h$ is lower semicontinuous. Let

$$k(t) = \inf\{g(t) : \ g \in C_{\mathbb{R}}(T), \ g \geq h\}$$

If we show that $k$ is continuous, then $k = \sup \mathcal{R}$; indeed, for all $f \in \mathcal{R}$ we have $f \leq h \leq k$ and if $f \leq g$ for all $f \in \mathcal{R}$ then $h \leq g$, and so $k \leq g$ by definition of supremum. Hence we focus the rest of the proof on showing that $k$ is continuous. By Corollary 1.8.13 we have that $k$ is upper semicontinuous. Fix $y \in \mathbb{R}$. Since $k$ is upper semicontinuous, $W_y = k^{-1}(-\infty, y)$ is open. As $T$ is extremally disconnected, $\overline{W_y}$ is clopen. This makes the function

$$g = y\,1_{\overline{W_y}} + (\|h\|_\infty + 1)\,1_{T\setminus\overline{W_y}}$$

continuous (by Exercise 1.8.15, and linear combinations of continuous are continuous). As $h \leq k$, $W_y \subset h^{-1}(-\infty, y]$; and so by the lower semicontinuity of $h$, $\overline{W_y} \subset h^{-1}(-\infty, y]$. This guarantees that $g \geq h$, and so $k \leq g$ by definition of $k$. In particular, $k \leq y$ on $\overline{W_y}$. So

$$\overline{k^{-1}(-\infty, y)} = \overline{W_y} \subset k^{-1}(-\infty, y].$$

All the above works for any $y \in \mathbb{R}$. So

$$\overline{k^{-1}(-\infty, y]} \subset \bigcap_n \overline{k^{-1}\Big(-\infty, y + \frac{1}{n}\Big)} \subset \bigcap_n k^{-1}\Big(-\infty, y + \frac{1}{n}\Big] = k^{-1}(-\infty, y].$$

This shows that $k^{-1}(-\infty, y]$ is closed for all $y \in \mathbb{R}$, and then $k$ is lower semicontinuous—and hence continuous—by Proposition 1.8.12. $\square$

For future reference, we have

***7.4.28. Proposition.***

*A metrizable Stonean space is finite.*

*Proof.* Let $M$ be Stonean and metric. We want to show that singletons are open. Fix $m \in M$. Consider the open set

$$V = \bigcup_n B_{1/2n}(m) \setminus \overline{B_{1/(2n+1)}(m)}.$$

We have that $m \in \overline{V}$, because $V$ contains elements arbitrarily close to it. As $\overline{V}$ is open, $m$ is an interior point. But by construction there exists a sequence in $M \setminus \overline{V}$ that converges to $m$, for the annuli $B_{1/(2n+1)}(m) \setminus \overline{B_{1/(2n+2)}(m)}$ are not in $V$. So $\{m\}$ is open. This shows that $M$ is discrete and compact, hence finite (Exercise 1.8.11). $\square$

A careful analysis of the proof of the Hahn–Banach theorem shows that the critical fact is that $\mathbb{C}$ is order-complete. Nachbin [Nac49] and Goodner [Goo50] noted this fact and showed a version of Hahn–Banach for $C(T)$ when

$T$ is extremally disconnected. The next three results are very strongly modelled on Lemma 5.7.3 and Theorems 5.7.4 and 5.7.5. We use the common convention that for functions $f, g$ the notation $f \leq g$ means $f(t) \leq g(t)$ for all $t$ in the domain of $f$. We interpret constants as constant functions.

**7.4.29. Lemma.**

*Let $T$ be a Stonean space, $V$ be a real vector space, $Z \subset V$ a subspace, and $v_0 \in V \setminus Z$. Suppose that $S : Z \to C_{\mathbb{R}}(T)$ is linear, and there exists a sublinear function $q : V \to \mathbb{R}$ such that $(Sx)(t) \leq q(x)$ for all $x \in Z$ and $t \in T$. Then there exists an extension $\tilde{S} : Z + \mathbb{R}v_0 \to C_{\mathbb{R}}(T)$ of $S$ with $\tilde{S}(x) \leq q(x)$ for all $x \in Z + \mathbb{R}v_0$.*

*Proof.* Exercise 7.4.11. □

**7.4.30. Proposition. (Sublinear Hahn–Banach Theorem for $C(T)$)**

*Let $T$ be a Stonean space, and $V$ be a real vector space with a sublinear function $q$. Let $W$ be a subspace of $V$, and $S : W \to C_{\mathbb{R}}(T)$ a linear map that satisfies $Sx \leq q(x)$ for all $x \in W$. Then there exists a linear map $\tilde{S} : V \to C_{\mathbb{R}}(T)$ such that $\tilde{S}|_W = S$ and $\tilde{S}x \leq q(x)$ for all $x \in V$.*

*Proof.* Exercise 7.4.12. □

**7.4.31. Theorem. (The Hahn–Banach Theorem for $C(T)$)**

*Let $T$ be a Stonean space, and $V$ be a complex vector space with a seminorm $q$. Let $W$ be a subspace of $V$ and $S : W \to C(T)$ a linear map satisfying $\|Sx\|_\infty \leq q(x)$ for all $x \in W$. Then there exists a linear map $\tilde{S} : V \to C(T)$ with $\tilde{S}|_W = S$ and $\|\tilde{S}x\|_\infty \leq q(x)$ for all $x \in V$.*

*The result holds with $C_{\mathbb{R}}(T)$ in place of $C(T)$, $S$ real-valued linear, and $q$ a real seminorm.*

*Proof.* Exercise 7.4.13. □

We will follow up on these ideas in Section 9.9.

We will now we build up the results we need to prove Theorem 7.4.36, which is a nice application of ideas in this section.

**7.4.32. Lemma.**

*Let $S,T$ be compact Hausdorff spaces with $T$ extremally disconnected and and $f : S \to T$ continuous and surjective. If $f|_{S_0}$ fails to be surjective for every proper closed subset $S_0 \subset S$, then $f$ is a homeomorphism.*

*Proof.* Suppose that $f(s_1) = f(s_2)$ with $s_1 \neq s_2$. Let $V_1, V_2$ be disjoint neighbourhoods of $s_1$ and $s_2$ respectively. From $V_1 \cap V_2 = \varnothing$ we get that $S = S\setminus(V_1 \cap V_2) = (S\setminus V_1)\cup(S\setminus V_2)$. Both $S\setminus V_1$ and $S\setminus V_2$ are compact, so $f(S\setminus V_1)$ and $f(S\setminus V_2)$ are compact (Exercise 1.8.37). Since $f$ is surjective, $f(S\setminus V_1)\cup f(S\setminus V_2) = T$, which implies

$$\big[T\setminus f(S\setminus V_1)\big]\cap\big[T\setminus f(S\setminus V_2)\big] = T\cap\big[f(S\setminus V_1)\cup f(S\setminus V_2)\big]^c = \varnothing.$$

By Exercise 1.8.45 (this is where we use the hypothesis) we get that $t = f(s_1) \subset f(V_1) \subset \overline{T\setminus f(S\setminus V_1)}$ and $t = f(s_2) \subset f(V_2) \subset \overline{T\setminus f(S\setminus V_2)}$. But this is a contradiction by Exercise 7.4.10. So $f$ is injective and thus bijective, and a homeomorphism by Exercise 1.8.38. $\square$

**7.4.33. Lemma.**

*Let $S,T$ be compact Hausdorff topological spaces and $f : S \to T$ continuous and surjective. Then there exists $S_0 \subset S$, compact, with $f(S_0) = T$ but such that $f$ is not surjective when restricted to any proper closed subset of $S_0$.*

*Proof.* Let $\mathcal{F} = \{S' \subset S : \text{ closed, and with } f(S') = T\}$, ordered by reverse inclusion. The family is nonempty because $S \in \mathcal{F}$. Note that any $S' \in \mathcal{F}$ is compact, because $S'$ is closed, and $S$ is compact Hausdorff (Lemma 1.8.16). Given a chain $\{S_j\}$ in $\mathcal{F}$, write $S' = \bigcap_j S_j$; for any $t \in T$ we have $S_j \cap f^{-1}(\{t\}) \neq \varnothing$, since $f(S_j) = T$. The sets $\{S_j \cap f^{-1}(\{t\})\}$ form a decreasing chain of compact subsets of $S$; given finitely many $S_{j_1} \supset \ldots \subset S_{j_m}$,

$$\bigcap_{k=1}^m S_j \cap f^{-1}(\{t\}) = \Big(\bigcap_{k=1}^m S_j\Big) \cap f^{-1}(\{t\}) = S_{j_m} \cap f^{-1}(\{t\}) \neq \varnothing.$$

Then $\displaystyle\bigcap_j S_j \cap f^{-1}(\{t\}) \neq \varnothing$ by Proposition 1.8.19. It follows that $f(S') = T$, and compact, so $S' \in \mathcal{F}$. Then Zorn's Lemma guarantees the existence of a maximal (properly, minimal, since we are using reverse inclusion) $S_0 \subset S$, which is a compact subset with $f(S_0) = T$ and such that no proper closed subset of $S_0$ is mapped to all of $T$. $\square$

***7.4.34. Proposition.***

*Let $T$ be a compact Hausdorff space. The following statements are equivalent:*

*(i)* *$T$ is extremally disconnected;*

*(ii)* *for every compact Hausdorff space $S$ and every continuous surjection $g : S \to T$ there exists a continuous right-inverse;*

*(iii)* *given compact Hausdorff spaces $R, S$ and continuous functions $g : S \to R$ and $f : T \to R$ with $g$ surjective, there exists continuous $h : T \to S$ such that the diagram*

$$\begin{tikzcd} T \arrow[d, "f"] \arrow[dr, dashed, "h"] & \\ R & S \arrow[l, "g"] \end{tikzcd} \tag{7.23}$$

*commutes.*

*Proof.* (i) $\Longrightarrow$ (ii) Let $g : S \to T$ be a continuous surjection. By Lemma 7.4.33 there exists $S_0 \subset S$, compact, with $g(S_0) = T$ and such that $g$ is not surjective when restricted to any proper compact subset of $S_0$. Then $g|_{S_0} : S_0 \to T$ is a homeomorphism by Lemma 7.4.32. Therefore $h = (g|_{S_0})^{-1}$ is a continuous right-inverse for $g$.

(ii) $\Longrightarrow$ (iii) Let $Z = \{(t, s) : \ t \in T, \ s \in S, \ f(t) = g(s)\}$ (this is known as the **fibre product**; the surjectivity of $g$ guarantees that $Z \neq \varnothing$). The continuity of $f, g$ and the fact that $T, S$ are compact Hausdorff guarantees that $Z$ is compact Hausdorff. Let $\pi_1 : Z \to T$ be the coordinate projection; this is a continuous surjection onto $T$ (because $g$ is surjective), so by hypothesis there exists a right-inverse $\rho : T \to Z$. Let $h = \pi_2 \circ \rho$. We know that $\pi_1(\rho(t)) = t$ for all $t \in T$. By definition of $Z$, $f = f \circ \pi_1 \circ \rho = g \circ \pi_2 \circ \rho = g \circ h$.

(iii) $\Longrightarrow$ (i) Fix $U \subset T$ open. Consider the disjoint union $S = \overline{U} \sqcup (T \setminus U)$, with the topology given by the disjoint union of the corresponding subspace topologies. Let $g : S \to T$ be given by $g(t) = t$ on each of both components. Then $g$ is surjective by construction (note that it is not injective, though this is not particularly relevant to us here), and it is also continuous for the preimages of open sets will yield open sets on each of the two components of the disjoint union. With $R = T$ and $f = \mathrm{id}$, the hypothesis gives us $h : T \to S$, continuous, with $g \circ h = \mathrm{id}$. As a subset of $S$, the closure $\overline{U}$ is open. Note that $g$ maps $\overline{U}$ onto $\overline{U}$ and $T \setminus \overline{U}$ into $T \setminus \overline{U}$. Then the continuity of $h$ gives us that for $\overline{U}$, as a subset of $T$,

$$\overline{U} = (g \circ h)^{-1}(\overline{U}) = h^{-1}(g^{-1}(\overline{U})) = h^{-1}(\overline{U}),$$

open, and hence $T$ is extremally disconnected. $\square$

**7.4.35. Corollary.**

*Let $T$ be a discrete topological space. Then $\beta T$ is extremally disconnected.*

*Proof.* Let $S$ be a compact Hausdorff space and $g : S \to \beta T$ continuous and surjective. For each $x \in \beta T$ choose $s_x \in g^{-1}(\beta T)$. So we get a function $f : T \to S$, given by $f(t) = s_{\delta(t)}$. As $T$ is discrete, $f$ is continuous. The universal property (7.18) of $\beta T$ gives us a continuous function $\alpha : \beta T \to S$ with $\alpha \circ \delta = f$. As $g \circ f = \delta$ by construction,

$$(g \circ \alpha) \circ \delta = g \circ f = \delta.$$

So the continuous function $g \circ \alpha : \beta T \to \beta T$ is the identity on $\delta(T)$, and hence by continuity $g \circ \alpha = \mathrm{id}_{\beta T}$. Then Proposition 7.4.34 implies that $\beta T$ is extremally disconnected. □

This next result appeared in [Gle58].

**7.4.36. Theorem. (Gleason)**

*Let $T$ be a compact Hausdorff space. Then there exists a Stonean space $X$ and a continuous surjection $\gamma : X \to T$ such that for any proper closed $X_0 \subset X$, $\gamma|_{X_0}$ is not surjective.*

*Proof.* Let $Y = T$ but considered with the discrete topology, and $X' = \beta Y$. Then $X'$ is extremally disconnected by Corollary 7.4.35. Let $f : Y \to T$ be the identity, which is certainly continuous since $Y$ is discrete. By the universal property (7.18) of $\beta Y$, there exists a continuous map $\gamma : \beta Y \to T$ that satisfies $\gamma(\delta(t)) = f(t) = t$ for all $t$. So $\gamma$ is a continuous surjection from $X' = \beta Y$ onto $T$. By Lemma 7.4.33 there exists $X \subset X'$, compact, such that $\gamma|_X$ is surjective but it fails to be surjective on any proper closed subset of $X$. If we consider

$$\begin{tikzcd} X' \arrow[d, "\gamma"] \arrow[dr, dashed, "h"] & \\ T & X \arrow[l, "\gamma"] \end{tikzcd}$$

since $\gamma$ is surjective on $X$, by (7.23) there exists $h : X' \to X$, continuous, with $\gamma \circ h = \gamma$. As $\gamma \circ h$ is surjective, $h(X')$ is compact, and $\gamma$ is not surjective on any proper closed subset of $X$, it follows that $h$ is surjective. If we consider $h|_X : X \to X$, it has to be surjective; for otherwise $\gamma : h(X) \to T$ would be surjective on the proper compact subset $h(X) \subset X$ (recall that $\gamma \circ h = \gamma$ and $\gamma|_X$ is surjective). If $h|_X \neq \mathrm{id}_X$, then by Lemma 1.8.14 there exists $E \subsetneq X$, compact, with $X = E \cup h(E)$. So if $x \in E$ then $\gamma(x) = \gamma(h(x)) \in \gamma(h(E))$, and if $x \in h(E)$ then $\gamma(x) \in \gamma(h(E))$. Then

$$T = \gamma(X) = \gamma(h(E)) = \gamma(E),$$

contradicting the minimality of $X$. Hence $h|_X = \mathrm{id}_X$.

Given $f : X \to R$ continuous and $g : S \to R$ continuous and surjective, we have

$$\begin{tikzcd} X \arrow[d, "f"] & X' \arrow[l, "h"'] \arrow[d, dashed, "\ell"] \\ R & S \arrow[l, "g"] \end{tikzcd}$$

and by $X'$ being extremally disconnected there exists continuous $\ell : X' \to S$ with $g \circ \ell = f \circ h$. But $h|_X = \mathrm{id}_X$, so $g \circ \ell|_X = f \circ h|_X = f$. Then $X$ is extremally disconnected by (7.23) via the function $\ell|_X : X \to S$. $\square$

**7.4.1. Exercises.**

(7.4.1) Show that $C_0(T)$ is complete.

(7.4.2) Show that if $T$ is locally compact Hausdorff, then $C_0(T)$ separates points: that is, given $s, t \in T$ with $s \neq t$, there exists $f \in C_0(T)$ with $f(s) \neq f(t)$.

(7.4.3) Show that if $X, Y$ are topological vector spaces, the following statements are equivalent:

(i) $C_{\mathbb{R}}(X)$ and $C_{\mathbb{R}}(Y)$ are isomorphic as rings;

(ii) $C_{\mathbb{R}}(X)$ and $C_{\mathbb{R}}(Y)$ are isomorphic as real algebras.

(7.4.4) Show that the function $d$ defined on (7.22) is indeed a metric, and that it induces the topology of $T$.

(7.4.5) Prove that a compact metric space is separable.

(7.4.6) Let $\{k_s\}_{s\in\mathbb{R}} \subset C_0(\mathbb{R})$ be the family of functions $k_s(t) = st(1+s^2t^2)^{-1}$, and $h(t) = (1+t)^{-1}$. Show that the algebra generated by $\{h\} \cup \{k_s\}$ is dense in $C_0(\mathbb{R})$.

(7.4.7) Let $S, T$ be locally compact Hausdorff spaces. Show that

$$\overline{\operatorname{span} C_0(S)C_0(T)} = C_0(S \times T),$$

where each product $fg$ is identified with the function $(s,t) \longmapsto f(s)g(t)$.

(7.4.8) Show that $c_0 = C_0(\mathbb{N})$.

(7.4.9) Show that the bounded set $\mathcal{R} = \{f \in C([0,1] : \ 0 \leq f \leq 1, \ f|_{[\frac12,1]} = 0\}$ in $C[0,1]$ admits no supremum.

(7.4.10) Let $T$ be an extremally disconnected topological space, and $U, V \subset T$ disjoint open subsets. Show that $\overline{U} \cap \overline{V} = \varnothing$.

(7.4.11) Prove Lemma 7.4.29 by modifying the proof of Lemma 5.7.3 appropriately.

(7.4.12) Prove Proposition 7.4.30 by modifying the proof of Theorem 5.7.4 appropriately.

(7.4.13) Prove Theorem 7.4.31 by adapting the proof of Theorem 5.7.5 appropriately.

(7.4.14) Let $T$ be a discrete topological space. Show that $\beta T$ is extremally disconnected without using Proposition 7.4.34. To show that an open set $V \subset \beta T$ has open closure, consider the function $f = 1_{V\cap\delta(T)} \in C(\delta(T))$.

## 7.5. Convexity

We recall terminology from Definition 4.3.3:

**7.5.1. Definition.**

*Given a subset $Y$ on a vector space $\mathcal{X}$, we denote by* $\operatorname{conv} Y$ *the* ***convex hull*** *of $Y$, i.e.*

$$\operatorname{conv} Y = \Big\{ \sum_{j=1}^n t_j y_j \,:\, n \in \mathbb{N}, y_1, \dots, y_n \in Y,\ t_1, \dots, t_n \in [0,1],\ \sum_{j=1}^n t_j = 1 \Big\}.$$

*It is the smallest convex subset of $\mathcal{X}$ that contains $Y$ (Exercise 7.5.3). The* ***closed convex hull*** *of $Y$ is $\overline{\operatorname{conv}} Y = \overline{\operatorname{conv} Y}$.*

*Given a convex subset $Y \subset \mathcal{X}$, a function $\phi : Y \to \mathbb{R}$ is* ***convex*** *if*

$$\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y), \qquad x, y \in Y,\ t \in [0,1].$$

A useful elementary result is that $\phi$ is convex if and only if for any convex combination $\sum_{j=1}^n t_j y_j$ with $y_1, \dots, y_n \in Y$ and $t_1, \dots, t_n \in [0,1]$ with $\sum_{j=1}^n t_j = 1$,

$$\phi\Big(\sum_{j=1}^n t_j y_j\Big) \leq \sum_{j=1}^n t_j\, \phi(y_j)$$

(Exercise 7.5.1).

In the real line we can say several specific things about convex functions. Note that the convex subsets of $\mathbb{R}$ are precisely the intervals (not necessarily finite, not necessarily open nor closed).

**7.5.2. Proposition.**

*Let $Y \subset \mathbb{R}$ convex, and $\phi : Y \to \mathbb{R}$ convex. The following statements are equivalent:*

*(i) $\phi$ is convex;*

*(ii) for any $y \in Y$, the function $x \longmapsto \dfrac{\phi(x) - \phi(y)}{x - y}$ is non-decreasing on $Y \setminus \{y\}$.*

*Proof.* (i) $\Longrightarrow$ (ii) Let $x_1, x_2 \in Y$ with $x_1 < x_2$. Assume first that $y < x_1$. Then we can write $x_1 = ty + (1-t)x_2$ for some $t \in (0,1)$, and

$$\begin{aligned} \frac{\phi(x_1) - \phi(y)}{x_1 - y} &= \frac{\phi(ty + (1-t)x_2) - \phi(y)}{(1-t)(x_2 - y)} \\ &\leq \frac{t\phi(y) + (1-t)\phi(x_2) - \phi(y)}{(1-t)(x_2 - y)} = \frac{\phi(x_2) - \phi(y)}{x_2 - y}. \end{aligned}$$

If $y > x_2$, we may reverse the roles of $x_1$ and $x_2$ above; as $x_2 - y < 0$, the inequality is reversed and we get again that

$$\frac{\phi(x_1) - \phi(y)}{x_1 - y} \leq \frac{\phi(x_2) - \phi(y)}{x_2 - y}.$$

And when $x_1 < y < x_2$ we can use the above, first $x_1$ as the pivoting point and then $x_2$:

$$\begin{aligned} \frac{\phi(x_1) - \phi(y)}{x_1 - y} &= \frac{\phi(y) - \phi(x_1)}{y - x_1} \leq \frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1} = \frac{\phi(x_1) - \phi(x_2)}{x_1 - x_2} \\ &\leq \frac{\phi(y) - \phi(x_2)}{y - x_2} = \frac{\phi(x_2) - \phi(y)}{x_2 - y}. \end{aligned}$$

(ii) $\Longrightarrow$ (i)Let $x, y \in Y$, $t \in (0,1]$ (there is nothing to prove if $t = 0$). Assume without loss of generality that $y < x$. Then $y \leq tx + (1-t)y \leq x$, so

$$\frac{\phi(tx + (1-t)y) - \phi(y)}{t(x - y)} \leq \frac{\phi(x) - \phi(y)}{x - y}.$$

Simplifying, since $t > 0$ and $x - y > 0$, we get $\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)$. ∎

**7.5.3. Proposition.**

*Let $Y \subset \mathbb{R}$ convex, and $\phi : Y \to \mathbb{R}$ convex. Then $\phi$ is continuous, differentiable almost everywhere, and*

$$\phi(x) = \sup\{ax + b : \; at + b \leq \phi(t) \text{ for all } t \in \mathbb{R}\}. \tag{7.24}$$

*Proof.* Fix $y \in Y$. By Proposition 7.5.2, the Newton quotients $\frac{\phi(x)-\phi(y)}{x-y}$ are monotone non-decreasing and bounded (with the possible exception of an endpoint) for $x < y$; this implies that the left derivative $\phi_l'(y)$ exists. Similarly, the right derivative $\phi_r'(y)$ also exists and $\phi_l' \leq \phi_r'$. In the same way we deduce that both functions $\phi_l'$ and $\phi_r'$ are monotone non-decreasing; in particular they are both continuous almost everywhere (Exercise 2.4.4). Given $\varepsilon > 0$ and for any $h$ with $0 < 2h < \varepsilon$,

$$\begin{aligned}\frac{\phi(y+h)-\phi(y)}{h} &= \frac{\phi(y+h)-\phi(y)}{(y+h)-y} = \frac{\phi(y)-\phi(y+h)}{y-(y+h)} \\ &\leq \frac{\phi(y+\varepsilon)-\phi(y+h)}{y+\varepsilon-(y+h)} = \frac{\phi(y+h)-\phi(y+\varepsilon)}{y+h-(y+\varepsilon)} \\ &\leq \frac{\phi(y+\varepsilon-h)-\phi(y+\varepsilon)}{y+\varepsilon-h-(y+\varepsilon)}.\end{aligned}$$

Taking the limit as $h \to 0^+$, the inequality $\phi_r'(y) \leq \phi_l'(y+\varepsilon)$ follows.

From $\phi_l'(y) \leq \phi_r'(y) \leq \phi_l'(r+\varepsilon)$, we deduce that at any point of continuity of $\phi_l'$ both side derivatives have to agree, making $\phi$ is differentiable; thus $\phi$ is differentiable almost everywhere. For the continuity of $\phi$, fix $y \in Y$ and $\varepsilon > 0$. For $h > 0$ and a fixed $h_0 > h$,

$$\phi(y+h)-\phi(y) = h\,\frac{\phi(y+h)-\phi(y)}{(y+h)-y} \leq h\,\frac{\phi(y+h_0)-\phi(y)}{y+h_0-y} = h\,\frac{\phi(y+h_0)-\phi(y)}{h_0}.$$

With the same argument,

$$\phi(y+h)-\phi(y) = h\,\frac{\phi(y+h)-\phi(y)}{(y+h)-y} \geq h\,\frac{\phi(y-h_0)-\phi(y)}{(y-h_0)-y} = -h\,\frac{\phi(y-h_0)-\phi(y)}{h_0}.$$

Taking limit as $h \to 0^+$ on both inequalities, we have shown that $\phi$ is continuous from the right. Continuity from the left can be shown using the exact same idea, and thus $\phi$ is continuous.

Finally, we look at (7.24). Let $s$ be the supremum in (7.24). If $at + b \leq \phi(t)$ for all $t$, we have in particular that $ax + b \leq \phi(x)$, so $\phi(x) \geq s$. Given $x \in Y$, we will show that there exists $a \in \mathbb{R}$ such that $a(t-x) + \phi(x) \leq \phi(t)$ for all $t$. This allows us to write $at + [-ax + \phi(x)] \leq \phi(t)$ for all $t$, which implies that $\phi(x) = ax + [-ax + \phi(x)] \leq s$, and so $s = \phi(x)$. To prove the inequality, we define $a = \phi_l'(x)$. Then, for $t > x$, Proposition 7.5.2 implies

$$a \leq \frac{\phi(t)-\phi(x)}{t-x},$$

which is equivalent to $a(t-x) + \phi(x) \leq \phi(t)$. And when $t < x$,

$$a \geq \frac{\phi(t)-\phi(x)}{t-x},$$

which again is equivalent to $a(t-x) + \phi(x) \leq \phi(t)$. $\square$

Given a convex function $\phi$, the line that is tangent to $\phi$ at $x$—which exists at all points where $\phi$ is differentiable, which is almost anywhere—is called the **support line** for $\phi$ at $x$. Such line will not exist at point where the derivative does not exist, like the point $x = 0$ for the convex function $\phi(x) = |x|$.

***7.5.4. Theorem. (Jensen's Inequality)***

*Let $(X, \Sigma, \mu)$ be a probability space, $g : X \to \mathbb{R}$ integrable, and $\phi : \mathbb{R} \to \mathbb{R}$ convex. Then*

$$\phi\Big(\int_X g\,d\mu\Big) \leq \int_X \phi \circ g\,d\mu.$$

*Proof.* We use the idea from the last paragraph of the proof of Proposition 7.5.3. Let $x = \displaystyle\int_X g\,d\mu$, and $a = \phi_l'(x)$. Then $\phi(t) \geq a(t - x) + \phi(x)$ for all $t$, and we have

$$\begin{aligned}\int_X \phi \circ g\,d\mu &\geq \int_X (a(g(t) - x) + \phi(x))\,d\mu(t)\\ &= a\int_X g\,d\mu - a\int_X x\,d\mu(t) + \int_X \phi(x)\,d\mu(t)\\ &= ax - ax + \phi(x) = \phi\Big(\int_X g\,d\mu\Big).\end{aligned}$$

□

*7.5.5. Remark .* An intuitive way of seeing Jensen's inequality arises when $g$ is simple. Indeed, if $g = \sum_k \beta_k 1_{E_k}$ then, since $\mu(E_1), \ldots, \mu(E_n)$ are convex coefficients,

$$\phi\Big(\int_X g\,d\mu\Big) = \phi\Big(\sum_k \mu(E_k)\,\beta_k\Big) \leq \sum_k \mu(E_k)\,\phi(\beta_k) = \int_X \phi \circ g\,d\mu.$$

Since $\phi$ is continuous, this can be made into a proof of Jensen's Inequality when one realizes that $\phi \circ g$ is again simple if $g$ is.

As seen in Exercise 5.1.5 the balls of $\ell^1(\mathbb{N})$ and $\ell^\infty(\mathbb{N})$ are intrinsically different from the balls in $\ell^p(\mathbb{N})$, $1 < p < \infty$; the latter are "round", while the former are not. The technical property that distinguishes these kinds of balls is **strict convexity**. Properly, a norm is strictly convex when any of the three equivalent properties below is satisfied.

**7.5.6. Proposition.**

*Let $\mathcal{X}$ be a normed space. The following statements are equivalent:*

*(i)* *if $x, y \in X$ are distinct and $\|x\| = \|y\| = 1$, then $\|x + y\| < 2$;*

*(ii)* *if $x, y \in X \setminus \{0\}$ and $\|x + y\| = \|x\| + \|y\|$, then there exists $\lambda > 0$ with $y = \lambda x$;*

*(iii)* *for each $\phi \in \mathcal{X}^*$ with $\|\phi\| = 1$, there exists at most one $x \in \mathcal{X}$ with $\|x\| = 1$ and such that $\phi(x) = 1$.*

*Proof.* Exercise 7.5.17. □

The norm $\|\cdot\|_p$ is strictly convex for $1 < p < \infty$, while $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not strictly convex (Exercise 7.5.18).

**7.5.7. Definition.**

*Given a convex set $K \subset \mathcal{X}$, a point $x \in K$ is an* ***extreme point*** *if $x = tx_1 + (1-t)x_2$ with $t \in [0,1]$, $x_1, x_2 \in K$ implies that $x_1 = x_2 = x$. The set of all extreme points in $K$ will be denoted by* $\operatorname{Ext} K$.

When $\mathcal{X}$ is locally convex, any extreme point of $K$ is necessarily a boundary point for $K$ (Exercise 7.5.2).

In basic examples in the plane, like a closed polygon or a closed disk $Y \subset \mathbb{R}^2$, one can see that $Y = \operatorname{conv} \operatorname{Ext} Y$. But this is not true in general, not even for closed sets in finite-dimension, as we will see momentarily. We will see in Theorem 7.5.11 that for compact convex sets, it is true that they are the closed convex hull of their extreme points.

**7.5.8. Proposition.**

*If $K \subset \mathcal{X}$ is convex and $x \in K$, then the following statements are equivalent:*

*(i)* $x \in \operatorname{Ext} K$;

*(ii)* *if $y, z \in K$ and $x = \frac{1}{2}(y + z)$, then $y = z = x$;*

*(iii)* *$K \setminus \{x\}$ is convex.*

*Proof.* Exercise 7.5.4. □

We mention some examples of convex sets and their extreme points:

(7.5.1) If $K$ is the closure of the interior of a regular polygon in $\mathbb{R}^2$, then $\operatorname{Ext} K$ is precisely the set of vertices of the polygon.
(7.5.2) If $K = \{(x,y) \in \mathbb{R}^2 : \ x^2 + y^2 \le 1\}$ (the closed unit disk), then $\operatorname{Ext} K = \{(x,y) \in \mathbb{R}^2 : \ x^2 + y^2 = 1\}$ (the unit circle).
(7.5.3) If $K = \{(x,y) \in \mathbb{R}^2 : \ y \ge 0\}$, then $K$ is closed, convex and $\operatorname{Ext} K = \varnothing$.
(7.5.4) If $K = \{(x,y) \in \mathbb{R}^2 : \ y > 0\} \cup \{(x_1, 0)\}$, then $K$ is convex and $\operatorname{Ext} K = \{(x_1, 0)\}$.
(7.5.5) If $K = \{x \in \ell^1(\mathbb{N}) : \ \|x\|_1 \le 1\}$, then $\operatorname{Ext} K = \{\lambda\delta_k : \ k \in \mathbb{N}, |\lambda| = 1\}$.
(7.5.6) If $1 < p < \infty$ and $K = \{x \in \ell^p(\mathbb{N}) : \ \|x\|_p \le 1\}$, then $\operatorname{Ext} K = \{x \in \ell^p(\mathbb{N}) : \ \|x\|_p = 1\}$.
(7.5.7) If $K = \{x \in \ell^\infty(\mathbb{N}) : \ \|x\|_\infty \le 1\}$, then $\operatorname{Ext} K = \{x \in \ell^\infty(\mathbb{N}) : \ |x(k)| = 1 \ \forall k \in \mathbb{N}\}$.
(7.5.8) If $K = \{f \in L^1[0,1] : \ \|f\|_1 \le 1\}$, then $\operatorname{Ext} K = \varnothing$.
(7.5.9) If $K = \{f \in C_0(\mathbb{R}) : \ \|f\|_\infty \le 1\}$, then $\operatorname{Ext} K = \varnothing$.
(7.5.10) If $T$ is locally compact and $K = \{\mu \in C_0(T)^* : \ \|\mu\| \le 1\}$, then $\operatorname{Ext} K = \{\lambda\,\delta_t : \ t \in T, \ |\lambda| = 1\}$.
(7.5.11) if $\mathbb{A} = \{f \in C(\overline{\mathbb{D}}) : \ f \ \text{ analytic in } \mathbb{D}\}$, is the **disk algebra**, then
$$\operatorname{Ext} \mathbb{A} = \Big\{f \in K : \ \int_0^{2\pi} (1 - \log|g(e^{it})|)\,dt = -\infty\Big\}.$$

We will consider (7.5.5)–(7.5.8) in some detail, and leave the rest as exercises; with the exception of (7.5.11), which requires some non-trivial complex analysis and appears in [**Hof62**, Theorem on page 138].

(7.5.5): We have $K = \{x \in \ell^1(\mathbb{N}) : \ \|x\|_1 \le 1\}$. If $\|x\|_1 < 1$, then $x = t(x/\|x\|_1) + (1-t)0$, where $t = \|x\|_1$, so $x \notin \operatorname{Ext} K$. If $x = e^{i\theta}\delta_k$ and $x = ty + (1-t)z$ with $\|y\|_1 \le 1$, $\|z\|_1 \le 1$, then
$$e^{i\theta} = x(k) = ty(k) + (1-t)z(k)$$
forces $|y(k)| = |z(k)| = 1$; since $\|y\|_1 \le 1$, this forces $y = e^{i\theta_1}\delta_k$, and similarly $z = e^{i\theta_2}\delta_k$. And since complex numbers with absolute value one are extreme in the unit disk in the complex plane, we necessarily have $\theta_1 = \theta_2 = \theta$, so $y = z = x$. Thus $x \in \operatorname{Ext} K$. Conversely, if $\|x\|_1 = 1$ and $x$ has at least two nonzero entries, then $|x(j)| < 1$ for all $j \in \mathbb{N}$. Say the two nonzero entries are the first two (it only simplifies notation). So $x(1) = re^{i\theta}$, $x(2) = se^{i\gamma}$, with $0 < r, s < 1$ and $r + s < 1$. Let $\varepsilon = \min\{r, s, 1-r, 1-s\}$. Define
$$y = (r+\varepsilon)e^{i\theta}\delta_1 + (s-\varepsilon)e^{i\gamma}\delta_2 + \sum_{j=3}^\infty x(j)\delta_j,$$
$$z = (r-\varepsilon)e^{i\theta}\delta_1 + (s+\varepsilon)e^{i\gamma}\delta_2 + \sum_{j=3}^\infty x(j)\delta_j.$$
Then $\|y\|_1 = \|z\|_1 = 1$, and $x = \frac{1}{2}\,y + \frac{1}{2}\,z$. So $x \notin \operatorname{Ext} K$.

(7.5.6): As with any unit ball, the fact that elements with norm less than 1 are not extreme follows the same proof as in the $p = 1$ case. If $\|x\|_p = 1$ and

$x = \frac{1}{2}\,y + \frac{1}{2}\,z$ with $\|y\|_p \le 1$ and $\|z\|_p \le 1$, then

$$1 = \|x\|_p = \frac{1}{2}\,\|y+z\|_p \le \frac{1}{2}\,\|y\|_p + \frac{1}{2}\,\|z\|_p \le 1,$$

so $\|y\|_p = \|z\|_p = 1$. Then $\|y+z\|_p = 2\|x\|_p = \|y\|_p + \|z\|_p$. As $1 < p < \infty$ we have that the norm is strictly convex, and then Proposition 7.5.6 gives us that $y = \lambda z$ with $\lambda > 0$, and then $\lambda = \|\lambda z\| = \|y\| = 1$; that is, $y = z$ and then $y = z = x$.

(7.5.7): If $\|x\|_\infty \le 1$ and there exists $k$ such that $|x(k)| < 1$, we can choose $c > 0$ such that $|x(k) \pm c| < 1$. Then $x = \frac{1}{2}\,(x + c\,\delta_k) + \frac{1}{2}\,(x - c\,\delta_k)$ and $x$ is not extreme. So any extreme point $x$ has $|x_k| = 1$ for all $k$. Any such $x$ is an extreme point, for if $x = \frac{1}{2}\,y + \frac{1}{2}\,z$ with $|x(k)| = 1$ for all $k$ and $\|y\|_\infty \le 1$, $\|z\|_\infty \le 1$, we have $x(k) = \frac{1}{2}\,y(k) + \frac{1}{2}\,z(k)$ for all $k$. As $|x(k)| = 1$, it follows that $|y(k)| = |z(k)| = 1$. We are now in the situation of Example (7.5.2), and so $y(k) = z(k) = x(k)$. Thus $y = z = x$ and $x$ is extreme.

(7.5.8): if $f \in L^1[0,1]$ and $\|f\|_1 = 1$, let $x \in [0,1]$ with $\int_0^x |f| = 1/2$. Then put $y = 2f\,1_{[0,x]}$, $z = 2f\,1_{[x,1]}$. So $\|y\|_1 = \|z\|_1 = 1$, and $f = \frac{1}{2}\,(y+z)$; so $f \notin \operatorname{Ext} K$.

As already mentioned, the difference between the cases for $\ell^1(\mathbb{N})$ and $\ell^p(\mathbb{N})$ above is mostly due to the fact that the unit ball in the case $p = 1$ is “square”, while it is “round” for $p > 1$. Recall Exercise 5.1.5. The variety of behaviour in the examples above is also somehow due to the fact that the unit balls in infinite dimension are not compact. This is not a very clear statement, but it is justified by the Krein Milman Theorem below (Theorem 7.5.11).

**7.5.9. Definition.**

*Given a convex $K \subset \mathcal{X}$, a nonempty subset $S$ of $K$ is said to be **extreme** if whenever $x, y \in K$, $0 < t < 1$ and $tx + (1-t)y \in S$, then $x \in S$, $y \in S$. A **face** is a convex extreme set. The set $\operatorname{Ext} K$ is precisely the union of the singleton extreme subsets $\mathcal{X}$.*

For a simple example, if $K \subset \mathbb{R}^2$ is a solid closed polygon, then the faces are $K$, each of the edges, and each vertex, while the extreme subsets are $K$, and any combination of unions of edges and vertices. For a closed disk, the extreme sets are the full disk, and any nonempty subset of the perimeter; and the faces are each point in the perimeter. For a closed cube in $\mathbb{R}^3$, the extreme subsets are the full cube, and any union of sides, edges, and vertices; while the faces are the full cube, each side, each edge, and each vertex.

**7.5.10. Lemma.**

*Let $\mathcal{X}$ be a topological vector space and $M \subset K \subset \mathcal{X}$, with $K$ compact and $M$ a compact extreme subset of $K$. Fix $\varphi \in \mathcal{X}^*$, and let $\mu = \max\{\operatorname{Re}\varphi(s) : \ s \in M\}$,*

$$M_\varphi = \{s \in M : \ \operatorname{Re}\varphi(s) = \mu\}.$$

*Then $M_\varphi$ is an extreme subset of $\mathcal{X}$.*

*Proof.* The maximum exists because $\varphi$ is continuous and $M$ is compact. Suppose that $ty + (1-t)z = x \in M_\varphi$, with $y, z \in K$, and $0 < t < 1$. Since $M$ is an extreme subset of $K$, we have $y, z \in M$. So $\operatorname{Re}\varphi(y) \leq \mu$, $\operatorname{Re}\varphi(z) \leq \mu$. But $x \in M_\varphi$, so $\operatorname{Re}\varphi(x) = \mu$ and this forces $\operatorname{Re}\varphi(y) = \operatorname{Re}\varphi(z) = \mu$. So $y, z \in M_\varphi$. □

**7.5.11. Theorem. (Krein–Milman)**

*Let $\mathcal{X}$ be a TVS such that $\mathcal{X}^*$ separates points, and let $K$ be a nonempty compact convex subset of $\mathcal{X}$. Then $K = \overline{\operatorname{conv}}\operatorname{Ext} K$.*

*Proof.* Let $\mathcal{P}$ be the set of all compact extreme subsets of $K$. Note that $K \in \mathcal{P}$, so $\mathcal{P}$ is nonempty. The first part of the proof consists of showing (7.25) below. This achieves two goals: the first one is to show that $\operatorname{Ext} K \neq \varnothing$, and the second one is the assertion (7.25) in itself; both facts are crucial to the second part of the proof.

Fix $S \in \mathcal{P}$. Let

$$\mathcal{Q} = \Big\{\{R_j\}_{j\in J} \subset \mathcal{P} : \ \text{ monotone decreasing, } R_j \subset S \ \text{ for all } j\Big\}.$$

Again this is nonempty, because $\{S\} \in \mathcal{Q}$. We can consider a partial order on $\mathcal{Q}$ given by inclusion (that is, $\{R_1, R_2, R_3\} \geq \{R_1, R_2\}$, etc.). Then we can use Zorn's Lemma to obtain a maximal totally ordered subcollection $\tilde{S}$ of $\mathcal{Q}$ (an upper bound for any chain is its union); the intersection $M$ of all of the members of $\tilde{S}$ is nonempty, because it is a collection of compact sets with the finite intersection property (Proposition 1.8.19). Moreover $M \in \mathcal{P}$, because an intersection of compact extreme subsets of $K$ is again a compact extreme subset of $K$.

No proper subset of $M$ can be in $\mathcal{P}$, because this would contradict the maximality of $\tilde{S}$. Fix $\varphi \in \mathcal{X}^*$, let $\mu = \max\{\operatorname{Re}\varphi(s) : \ s \in M\}$ (recall that $M$ is compact and $\varphi$ is continuous), and let

$$M_\varphi = \{s \in M : \ \operatorname{Re}\varphi(s) = \mu\}.$$

By Lemma 7.5.10, $M_\varphi \in \mathcal{P}$. This implies that $M = M_\varphi$, because $M$ is minimal in $\mathcal{P}$. So $\operatorname{Re}\varphi$ is constant on $M$; as the real parts of the functionals in $\mathcal{X}^*$ separate points (Exercise 5.7.13), $M$ has to consist of a single point, which is in $\operatorname{Ext} K$ since $M$ is extreme.

So far we have proven that

$$S \cap \operatorname{Ext} K \neq \varnothing, \qquad S \in \mathcal{P}. \tag{7.25}$$

With (7.25) established, we can now show the equality $K = \overline{\operatorname{conv}} \operatorname{Ext} K$. Using that $K$ is convex and compact and that $\operatorname{Ext} K \subset K$, it follows readily that $\overline{\operatorname{conv}} \operatorname{Ext} K \subset K$. If this inclusion were proper, we would have some $x_0 \in K \setminus \overline{\operatorname{conv}} \operatorname{Ext} K$. Now we would like to apply the Geometric Hahn–Banach theorem, but we find the problem that $\mathcal{X}$ was not assumed locally convex. Recall, however, that $\mathcal{X}$ is locally convex when considered with its weak topology (Remark 7.1.7), and that in such situation the dual of $\mathcal{X}$ stays the same (Proposition 7.1.5). As $\overline{\operatorname{conv}} \operatorname{Ext} K$ is compact—closed subset of compact is always compact—it is also weakly compact and so it is weakly closed—here we use that the weak topology is Hausdorff when $\mathcal{X}^*$ separates points, Lemma 7.1.2. So we can apply Theorem 5.7.18 (second form of Geometric Hahn–Banach) to $(\mathcal{X}, \sigma(\mathcal{X}, \mathcal{X}^*))$ to obtain $\varphi \in \mathcal{X}^*$ such that

$$\operatorname{Re} \varphi(x) < \operatorname{Re} \varphi(x_0), \qquad x \in \overline{\operatorname{conv}} \operatorname{Ext} K. \tag{7.26}$$

By Lemma 7.5.10, now applied to $K$, the set $K_\varphi$ is in $\mathcal{P}$. But the inequality (7.26) guarantees that $K_\varphi \cap \operatorname{Ext} K = \varnothing$, contradicting (7.25). So $\overline{\operatorname{conv}} \operatorname{Ext} K = K$. □

In finite dimension one can prove Krein–Milman without much machinery:

***7.5.12. Theorem. (Carathéodory)***

*Let $K \subset \mathbb{R}^n$ be compact and convex. Then $K = \operatorname{conv} \operatorname{Ext} K$, and every element of $K$ can be written as a convex combination of at most $n+1$ extreme points.*

*Proof.* We proceed by induction on $n$. When $n = 0$ we have a single point which is then extreme. If $K \subset \mathbb{R}^{n+1}$ is compact and $x \in K$, we do the following. Let $y$ be some extreme point of $K$ (which exists by Exercise 7.5.6). Consider the segment

$$S = \{y + t(x-y) : \ t \geq 0\}$$

This is the line segment that starts at $y$, passes through $x$, and keeps going. $S$ is closed, so $S \cap K$ is compact. Let $t_m = \max\{t \geq 1 : \ y + t(x-y) \in K\}$. The convexity and compactness of $K$ guarantee that $t_m$ is well-defined, and $t_m \in \partial K$. Let $H$ be a supporting hyperplane for $K$ at $x_m = y + t_m(x-y)$ (see Exercise 7.5.8). Then $H \cap K$ is a compact convex set, and $\operatorname{Ext}(H \cap K) \subset \operatorname{Ext} K$ (Exercise 7.5.9). By translating to the origin, we may see $H \cap K$ as a convex subset of an $n$-dimensional space. The inductive hypothesis then implies that

$x_m = \sum_{j=1}^n t_j z_j$ for convex coefficients $t_j$ and $z_1, \ldots, z_n \in \operatorname{Ext} H \cap K \subset \operatorname{Ext} K$. If $t_m = 1$, then $x_m = x$ and we are done. Otherwise $t_m > 1$, and

$$x = \Big(1 - \frac{1}{t_m}\Big) y + \frac{1}{t_m} x_m$$

is a convex combination of $n+1$ extreme points. □

When one does not care about extreme points, a more general version of Theorem 7.5.12 can be proven; we refer to Exercise 7.5.26.

The Krein–Milman Theorem is a very useful tool for existence results, as it guarantees the existence of the extreme points. Here is a non-existence application, that has been mentioned before:

**7.5.13. Proposition.**

*Neither of the Banach spaces $L^1[0,1]$ and $C_0(T)$ (for any locally compact non-compact space $T$) is isometrically isomorphic to the dual of any Banach space.*

*Proof.* Suppose that $\mathcal{X}^* = L^1[0,1]$ or $\mathcal{X}^* = C_0(T)$ for some Banach space $\mathcal{X}$. We apply Krein–Milman (Theorem 7.5.11) to the closed unit ball of $\mathcal{X}^*$ with the weak$^*$-topology. This implies that the closed unit ball in $\mathcal{X}^*$ has extreme points; but that is not the case, as proven on Examples (7.5.8) and (7.5.9). □

It turns out that for $L^1[0,1]$ a stronger statement can be proven; namely, that $L^1[0,1]$ cannot be embedded in any separable dual space. We refer to [**AK06**, Theorem 6.3.7].

**7.5.14. Lemma.**

*Let $\mathcal{X}$ be a TVS, and let $K_1, \ldots, K_m$ be convex and compact. Then $\operatorname{conv}(\bigcup_{r=1}^m K_r)$ is compact.*

*Proof.* If any $K_r$ is empty it contributes nothing, so we will assume $K_r \neq \varnothing$ for all $r$. An element $x \in \operatorname{conv}(\displaystyle\bigcup_{r=1}^m K_r)$ is of the form (after possibly reordering to group elements from the same $K_r$)

$$x = \sum_{r=1}^m \sum_{j=1}^{s_r} t_{r,j} x_{r,j},$$

where $x_{r,j} \in K_r$ for all $r, j$ and $t_{r,j} \geq 0$ for all $r, j$ and $\sum_{r,j} t_{r,j} = 1$. Let

$$\alpha_r = \sum_{j=1}^{s_r} t_{r,j}, \qquad\qquad x_r = \frac{1}{\alpha_r} \sum_{j=1}^{s_r} t_{r,j} x_{r,j}, .$$

unless $\alpha_r = 0$ in which case we take $x_r$ to be any element of $K_r$. We have $\alpha_r \geq 0$ for all $r$ and $\sum_r \alpha_r = 1$. Then

$$x = \sum_{r=1}^{m} \sum_{j=1}^{s_r} t_{r,j} x_{r,j} = \sum_{r=1}^{m} \alpha_r \frac{1}{\alpha_r} \sum_{j=1}^{s_r} t_{r,j} x_{r,j} = \sum_{r=1}^{m} \alpha_r x_r.$$

In words, any element of $\operatorname{conv}(\bigcup_{r=1}^m K_r)$ can be written as a convex combination with precisely one element from each set. Now let $T \subset \mathbb{R}^m$ be the set

$$T = \Big\{(t_1, \dots, t_m) : \ t_j \geq 0 \ \text{ for all } j \text{ and } \sum_j t_j = 1\Big\}.$$

Then $T$ is a closed and bounded subset of $\mathbb{R}^n$ and hence compact. So the set $T \times K_1 \times \cdots \times K_m$ is compact. The function $\alpha : T \times K_1 \times \cdots \times K_m \to \mathcal{X}$ given by

$$\alpha((t_1, \dots, t_m), x_1, \dots, x_m) = \sum_{j=1}^{m} t_j x_j$$

is continuous by definition of TVS. Hence

$$\operatorname{conv}(\bigcup_{r=1}^{m} K_r) = \alpha(T \times K_1 \times \cdots \times K_m)$$

is compact by Exercise 1.8.37. $\square$

The proof of Lemma 7.5.14 is neat, but it is not something that one would come up with on a first try. It is not too hard, though, to show Lemma 7.5.14 in a more straightforward way (Exercise 7.5.23). On the topic of the proof of Lemma 7.5.14, convexity of $K_j$ is not mentioned in the proof; is it used? Where? (Exercise 7.5.21).

*7.5.15. Example.* Without closure, the convex hull of a compact set may fail to be compact. For instance let $\mathcal{X} = \ell^1(\mathbb{N})$ and define $x_n = \frac{1}{n} e_n$, $n \in \mathbb{N}$. Let $K = \{x_n\} \cup \{0\}$. Then $K$ is compact, but

$$\operatorname{conv} K = \Big\{ \sum_{k=1}^{m} t_k x_k : \ \sum_{k=1}^{m} t_k = 1, \ t_k \geq 0 \ \text{ for all } k \Big\}$$

is not closed as it fails to have $\sum_k 2^{-k} x_k \in \overline{\operatorname{conv}} K$.

*7.5.16. Example.* In general if $K$ is compact but not convex, even with closure $\overline{\operatorname{conv}}K$ may fail to be compact. When that happens, it might be the case that $\overline{\operatorname{conv}}K$ has extreme points which are not in $K$. To see an example of this, let $\mathcal{X} = C[0,1]$ with the supremum norm, and let $K = \{\delta_t : \ t \in [0,1]\} \subset \mathcal{X}^*$, the characters, where we consider the weak*-topology. Since $t \mapsto \delta_t$ is continuous, $K$ is the image of the compact set $[0,1]$ under this map, so it is compact. Let $\varphi \in \mathcal{X}^*$ be given by $\varphi(f) = \int_0^1 f$. For any $f \in C[0,1]$,

$$\varphi(f) = \int_0^1 f = \lim_{n\to\infty} \sum_{k=1}^n f\left(\frac{k}{n}\right)\frac{1}{n} = \lim_{n\to\infty} \sum_{k=1}^n \frac{1}{n}\,\delta_{k/n}(f), \tag{7.27}$$

so $\varphi \in \overline{\operatorname{conv}}K$. Any $\psi \in \overline{\operatorname{conv}}K$ satisfies $\psi(f) \geq 0$ for all $f \geq 0$ and $\psi(1) = 1$. Conversely, let $\psi \in \mathcal{X}^*$ with $\psi(f) \geq 0$ for all $f \geq 0$ and $\psi(1) = 1$. By Riesz–Markov Theorem 2.9.2 there exists a measure $\mu$ such that $\psi(f) = \int_{[0,1]} f\,d\mu$. From $\psi(1) = 1$ we get $\mu([0,1]) = 1$. Define intervals $I_{k,n} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$, and let $t_k = \frac{k}{2^n}$. Then, for any $f \in \mathcal{X}$,

$$\left|\psi(f) - \sum_{k=0}^{2^n-1} \mu(I_{k,n})\,\delta_{t_k}(f)\right| \leq \sum_{k=1}^{2^n-1} \int_{I_{k,n}} |f(t) - f(t_k)|\,d\mu(t).$$

As $f$ is uniformly continuous, we get that

$$\psi = \lim_{n\to\infty} \sum_{k=0}^{2^n-1} \mu(I_{k,n})\,\delta_{t_k} \in \overline{\operatorname{conv}}K.$$

Thus

$$\overline{\operatorname{conv}}K = \{\psi \in \mathcal{X}^* : \ \psi(f) \geq 0 \ \text{ for all } f \geq 0,\ \psi(1) = 1\}.$$

Now let

$$\mathcal{M} = \operatorname{span}\Big\{\{\varphi\} \cup K\Big\},$$

still with the weak*-topology of $\mathcal{X}$. We have $K \subset \mathcal{M}$. By (7.27), $\varphi \in \mathcal{M} \cap \overline{\operatorname{conv}}K$, which is the closed convex hull of $K$ in $\mathcal{M}$. Suppose that $\varphi = \alpha\,\psi_1 + (1-\alpha)\,\psi_2$, for $\psi_1, \psi_2 \in \mathcal{M} \cap \overline{\operatorname{conv}}K$. As each of $\psi_1, \psi_2$ is a linear combination of $\varphi$ and finitely many point evaluations, this means

$$\varphi = \alpha_0\varphi + \sum_{j=1}^m \alpha_j\delta_{t_j},$$

with $t_1, \ldots, t_m \in [0,1]$ and $\alpha_0, \ldots, \alpha_m \in \mathbb{C}$. This we may rewrite, if $\alpha_0 \neq 1$, as

$$\varphi = \tfrac{1}{1-\alpha_0} \sum_{j=1}^m \alpha_j\delta_{t_j}.$$

Evaluating at any nonzero $g \in \mathcal{X}$ with $g \geq 0$ and $g(t_j) = 0$ for all $j = 1, \ldots, m$, we get a contradiction, showing that $\alpha_0 = 1$. Then, from $\sum_{j=1}^m \alpha_j\delta_{t_j} = 0$ we get, evaluating at each $t_j$, that $\alpha_1, \ldots, \alpha_m = 0$. This means that $\psi_1 = \psi_2 = \varphi$ and so $\varphi \in \operatorname{Ext}(\mathcal{M} \cap \overline{\operatorname{conv}} K)$, while $\varphi \notin K$. By Theorem 7.5.18, we have constructed an example of a situation where $\mathcal{M}$ is a LCS, $K \subset \mathcal{M}$ is compact, and $\overline{\operatorname{conv}} K$ is not compact.

*7.5.17. Remark* . When $\mathcal{X}$ is normed, though, if $K \subset \mathcal{X}$ is compact then $\overline{\operatorname{conv}} K$ is compact, too. Indeed, fix $\varepsilon > 0$. Since $K$ is compact, there exist $x_1, \dots, x_m \in K$ such that $K \subset \bigcup_{k=1}^m B_\varepsilon(x_k)$. We can write this as

$$K \subset F + B_\varepsilon(0),$$

where $F = \{x_1, \dots, x_m\}$. It follows that

$$\operatorname{conv} K \subset B_\varepsilon(0) + \operatorname{conv} F,$$

since $B_\varepsilon(0)$ is already convex. As $F$ is finite, $\operatorname{conv} F$ is compact (since $\overline{\operatorname{conv}} F$ is compact as it is a closed and bounded subset of a finite-dimensional space, and $\overline{\operatorname{conv}} F = \operatorname{conv} F$ by Lemma 7.5.14 since a normed space is Hausdorff). So there exist $f_1, \dots, f_r \in \operatorname{conv} F$ such that $\operatorname{conv} F \subset \bigcup_{j=1}^r B_\varepsilon(f_j)$. Then,

$$\operatorname{conv} K \subset \bigcup_{j=1}^r B_{2\varepsilon}(f_j),$$

and then

$$\overline{\operatorname{conv}} K \subset \bigcup_{j=1}^r B_{3\varepsilon}(f_j).$$

By Lemma 1.8.25, $\overline{\operatorname{conv}} K$ is compact.

The pathology about extreme points in Example 7.5.16 cannot happen when $\overline{\operatorname{conv}} K$ is compact. We have

***7.5.18. Theorem. (Milman)***

*If $\mathcal{X}$ is a LCS, $K \subset \mathcal{X}$ is compact and $\overline{\operatorname{conv}} K$ is compact, then*

$$\operatorname{Ext} \overline{\operatorname{conv}} K \subset K.$$

*Proof.* Let $p \in \operatorname{Ext} \overline{\operatorname{conv}} K \setminus K$. By Theorem 5.4.6 there exists $V \subset \mathcal{X}$, open, convex, and balanced with $0 \in V$, such that

$$(p + V) \cap (K + V) = \varnothing. \tag{7.28}$$

As $K$ is compact, there exist $k_1, \dots, k_m \in K$ with $K \subset \bigcup_{j=1}^m (k_j + V)$.

Let $R_j = \overline{\operatorname{conv}}(K \cap (k_j + V))$. As $k_j + \overline{V}$ is closed and convex, we have $R_j \subset k_j + \overline{V}$. Since $\overline{\operatorname{conv}} K$ is compact and $R_j \subset \overline{\operatorname{conv}} K$, we have that $R_j$ is compact. By Lemma 7.5.14, $\operatorname{conv}(\bigcup_{j=1}^m R_j)$ is closed. Then

$$\operatorname{conv}\Big(\bigcup_{j=1}^m R_j\Big) \subset \overline{\operatorname{conv}} K \subset \overline{\operatorname{conv}}\Big(\bigcup_{j=1}^m R_j\Big) = \operatorname{conv}\Big(\bigcup_{j=1}^m R_j\Big).$$

As $p \in \overline{\operatorname{conv}} K = \operatorname{conv}\bigcup_{j=1}^m R_j$, we can write $p = \sum_{j=1}^m t_j p_j$, with $\{t_j\}$ convex coefficients and $p_j \in R_j \subset k_j + \overline{V} \subset K + \overline{V}$. As $p$ is extreme in $\overline{\operatorname{conv}} K$ and $p_j \in R_j \subset \overline{\operatorname{conv}} K$ for all $j$, necessarily $p_1 = \cdots = p_m = p$. Thus $p \in K + \overline{V}$, contradicting the separation (7.28). ☐

### 7.5.1. Exercises.

(7.5.1) Show that $\phi : \mathcal{Y} \to \mathbb{R}$ is convex if and only if

$$\phi\Big(\sum_{j=1}^n t_j y_j\Big) \le \sum_{j=1}^n t_j\, \phi(y_j) \tag{7.29}$$

for all $y_1, \ldots, y_n \in \mathcal{Y}$ and all $t_1, \ldots, t_n \in [0,1]$ with $\sum_j t_j = 1$.

(7.5.2) Show that if $\mathcal{X}$ locally is convex, then any extreme point of the convex set $K \subset \mathcal{X}$ is a boundary point of $K$.

(7.5.3) Let $\mathcal{X}$ be a real or complex vector space and $Y \subset \mathcal{X}$. Show that $\operatorname{conv} Y$ is the smallest convex subset of $\mathcal{X}$ that contains $Y$.

(7.5.4) Prove Proposition 7.5.8.

(7.5.5) Let $C \subset \mathbb{R}^n$ be convex, and $w \in \mathbb{R}^n \setminus C$. Show that there exists $v \in \mathbb{R}^n$ with $\|v\| = 1$ and $\alpha \in \mathbb{R}$ such that $\langle w, v\rangle \le \alpha$ and $\langle z, v\rangle \ge \alpha$ for all $z \in C$. This can be done via Hahn–Banach, but an elementary proof is desired.

(7.5.6) Let $K \subset \mathbb{R}^n$ be compact and convex. Show that $K$ has an extreme point. *(Hint: choose two points that realize the diameter, and show that they are extreme).*

(7.5.7) A **hyperplane** is a subset $H \subset \mathbb{R}$ of the form $z + V$, where $V \subset \mathbb{R}^n$ is subspace of dimension $n-1$. Show that $H \subset \mathbb{R}^n$ is a hyperplane if and only if there exist $\alpha \in \mathbb{R}$ and $v \in \mathbb{R}^n$ such that $H = \{x \in \mathbb{R}^n : \langle x, v\rangle = \alpha\}$.

(7.5.8) Let $C \subset \mathbb{R}^n$ be convex. A **supporting hyperplane** for $C$ is a hyperplane $H \subset \mathbb{R}^n$ such that

(i) $C$ is in one of the two half-spaces determined by $H$; namely, there exists $\alpha \in \mathbb{R}$ such that $\langle x, y\rangle \ge \alpha$ for all $x \in C$ and $y \in H$;

(ii) $H \cap \partial C \neq \varnothing$.

Show that for each $x \in \partial C$ there exists a supporting hyperplane $H$ for $C$ such that $x \in H$.

(7.5.9) Let $C \subset \mathbb{R}^n$ be convex, $x \in \partial C$ and $H$ a supporting hyperplane for $C$ at $x$. Show that $H$ is convex and that $\operatorname{Ext}(H \cap C) \subset \operatorname{Ext} C$.

(7.5.10) In Example (7.5.1), show that $K$ is convex and establish $\operatorname{Ext} K$.
(7.5.11) In Example (7.5.2), show that $K$ is convex and establish $\operatorname{Ext} K$.
(7.5.12) In Example (7.5.3), show that $K$ is closed, convex and establish $\operatorname{Ext} K$.
(7.5.13) In Example (7.5.4), show that $K$ is convex and establish $\operatorname{Ext} K$.
(7.5.14) Show that, in the complex plane $\operatorname{Ext} \overline{\mathbb{D}} = \mathbb{T}$.
(7.5.15) In Example (7.5.9), show that $K$ is convex and establish $\operatorname{Ext} K$.
(7.5.16) In Example (7.5.10), show that $K$ is convex and establish $\operatorname{Ext} K$.
(7.5.17) Prove Proposition 7.5.6.
(7.5.18) Show that $\|\cdot\|_p$ is strictly convex for $1 < p < \infty$, while $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not strictly convex.
(7.5.19) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $V \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ an isometry.

(i) Show that if $V$ is surjective, then $V$ maps the set $\operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)}$ onto $\operatorname{Ext} \overline{B_1^{\mathcal{Y}}(0)}$.

(ii) Show that if $V$ is not surjective, the above may fail, i.e. construct an example of a linear isometry that maps an extreme point to a non-extreme point.

(iii) *(This one may be a little harder because we want the same domain and codomain; but examples exist that are not convoluted)* Find a Banach space $\mathcal{X}$ with $\operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)} \neq \varnothing$, an isometry $V \in \mathcal{B}(\mathcal{X})$, and $e \in \operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)}$ such that $Ve \notin \operatorname{Ext} \overline{B_1^{\mathcal{X}}(0)}$.

(7.5.20) Prove Exercise 5.6.5 using convexity ideas. That is, show that $c$ and $c_0$ are not isometrically isomorphic as Banach spaces (*Hint: show that the unit ball of $c_0$ has no extreme points, while the unit ball of $c$ does, and use Exercise 7.5.19*)
(7.5.21) Is convexity of $K_j$ used in the proof of Lemma 7.5.14? Where?
(7.5.22) Let $\mathcal{X}$ be a TVS, $K \subset \mathcal{X}$, and $x_1, x_2, \ldots \in K$. Show that $\displaystyle\sum_{k=1}^\infty 2^{-k} x_k \in \overline{\operatorname{conv}} K$.
(7.5.23) Prove Lemma 7.5.14 in a straightforward way, that is without using the fact that a continuous image of compact is compact.
(7.5.24) For Hilbert spaces, we proved in Lemma 4.3.4 that given a closed convex set $K$ and $x \notin K$, the distance between $x$ and $K$ is achieved, and it is achieved at a unique point. The same is not true for an arbitrary Banach space. Let $\mathcal{X} = C[0,1]$, with the supremum norm, and let

$$K = \left\{ g \in C[0,2] : \int_0^1 g - \int_1^2 g = 1 \right\}$$

Show that $K$ is closed and convex, that $\operatorname{dist}(0, K) = \frac{1}{2}$, and that $\|g\| > \frac{1}{2}$ for all $g \in K$.

(7.5.25) Given $X \subset \mathbb{R}^n$, the **cone** generated by $X$ is the set

$$\operatorname{cone} X = \Big\{ \sum_{k=1}^{m} \alpha_k x_k : \ m \in \mathbb{N},\ x_1, \dots, x_m \in X,\ \alpha_1, \dots, \alpha_m \geq 0 \Big\}.$$

Let $x_1, \dots, x_m \in \mathbb{R}^n$ and $x \in \operatorname{cone}\{x_1, \dots, x_m\}$. Show that there exist $\alpha_1, \dots, \alpha_m \in [0, \infty)$, with at most $n$ of them nonzero, such that $x = \sum_k \alpha_k x_k$.

(7.5.26) Let $X \subset \mathbb{R}^n$.

(i) Show that if $x \in \operatorname{cone} X$ then $x$ can be written as a positive linear combination of at most $n$ points in $X$.

(ii) Show that if $x \in \operatorname{conv} X$ then $x$ can be written as a convex combination of at most $n+1$ points in $X$.

(7.5.27) Let $X \subset \mathbb{R}^n$ be compact. Show that $\operatorname{conv} X$ is compact.

## 7.6. The Cantor Space

We will now revisit the Cantor set from Section 2.2, and show that it is not just a curiosity nor purely a source of counterexamples, but an important topological object in itself.

**7.6.1. Definition.**

*We write $2^{\mathbb{N}} = \{0,1\}^{\mathbb{N}}$, considered as a topological space with pointwise convergence. This space is compact by Tychonoff's Theorem. It is often called the* ***Cantor Space****.*

The reason why $2^{\mathbb{N}}$ is called the Cantor Space is the following:

***7.6.2. Proposition.***

*There is a homeomorphism between $\mathcal{C}$ and $2^{\mathbb{N}}$.*

*Proof.* Define $\lambda : 2^{\mathbb{N}} \to \mathcal{C}$ by $\lambda((c_n)) = 2\sum_{n=1}^{\infty} \frac{c_n}{3^n}$. Proposition 2.2.1 guarantees that $\lambda$ takes values in $\mathcal{C}$. We need to check that $\lambda$ is bijective, and that it is continuous with continuous inverse.

To see that $\lambda$ is bijective, we simply write $\lambda^{-1}$ explicitly: since under the convention to write elements in $\mathcal{C}$ not ending in 1 the representation of each number is unique, the map

$$\lambda^{-1} : 2\sum_{n=1}^{\infty} \frac{c_n}{3^n} \longmapsto \{c_n\}$$

is well-defined. Continuity: suppose that $\{c_n^k\} \underset{k}{\to} \{c_n\}$ pointwise (as this is the convergence in $2^{\mathbb{N}}$). Fix $\varepsilon > 0$. There exists $n_0$ such that $\sum_{n>n_0} 3^{-n} < \varepsilon$. Choose $k_0$ such that $|c_n^k - c_n| < \min\{\varepsilon, 1\}$, for $n = 1, \ldots, n_0$ if $k > k_0$ (this means that $c_n^k = c_n$). Then

$$|\lambda(\{c_n^k\}) - \lambda(\{c_n\})| \leq 2\sum_{n=1}^{n_0} \frac{|c_n^k - c_n|}{3^n} + 2\sum_{n>n_0} \frac{2}{3^n} = 2\sum_{n>n_0} \frac{2}{3^n} \leq 4\varepsilon.$$

So $\lambda$ is a continuous bijection with compact domain, and is thus a homeomorphism. $\square$

A **continuous retract** of a subspace $S$ of the topological space $T$ is a continuous map $\gamma : T \to S$ with $\gamma(s) = s$ for all $s \in S$.

***7.6.3. Proposition.***

*Let $S \subset \mathcal{C}$ be closed. Then $S$ admits a continuous retract.*

*Proof.* Let $\gamma\Big(\sum_{n=1}^{\infty} \frac{a_n}{3^n}\Big) = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$, where $b_n$ is defined inductively as follows:

$$b_{n+1} = \begin{cases} a_{n+1}, & \text{if } \exists \displaystyle\sum_{k=1}^{\infty} \frac{c_k}{3^k} \in S \text{ with } (c_1, \ldots, c_{n+1}) = (a_1, \ldots, a_{n+1}) \\ 2 - a_{n+1}, & \text{otherwise} \end{cases}$$

By definition, $\gamma(s) = s$ for all $s \in S$. Let $s = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \in \mathcal{C} \setminus S$. If $a_1, \ldots, a_n$ are the first $n$ digits of some element in $S$ and there is no element in $S$ with first digits $a_1, \ldots, a_{n+1}$, there is necessarily an element in $S$ with first digits $a_1, \ldots, a_n, 2 - a_{n+1}$; since $2 - a_{n+1}$ is the only other possible digit. This means that $\{b_n\}$ was constructed in such a way that, given $n$, there exists an element in $s' \in S$ with first digits $b_1, \ldots, b_n$; and so

$$|s' - \gamma(s)| = \left| s' - \sum_{k=1}^{\infty} \frac{b_k}{3^k} \right| \leq \sum_{k=n+1}^{\infty} \frac{2}{3^k} = \frac{1}{3^n}$$

We have shown that $\gamma(s)$ is a limit of points in $S$; as $S$ is closed, $\gamma(s) \in S$. The continuity of $\gamma$ is clear: if $|s - t| < 3^{-n}$, then $s$ and $t$ coincide in their first $n$ entries of their ternary representations; the same will be true of $\gamma(s)$ and $\gamma(t)$, so $|\gamma(s) - \gamma(t)| < 3^{-n}$. □

As we did with the case of $2^{\mathbb{N}}$, for any topological space $X$ we consider $X^{\mathbb{N}} = \{g : \mathbb{N} \to X\}$ with the pointwise convergence topology, which is precisely the product topology. Let us consider what this gives us in the case of $(2^{\mathbb{N}})^{\mathbb{N}}$. The elements of $(2^{\mathbb{N}})^{\mathbb{N}}$ are sequences $\{g_m\}$ where $g_m \in 2^{\mathbb{N}}$ for all $m$; so $\{p_\ell\} \subset (2^{\mathbb{N}})^{\mathbb{N}}$ converges to 0 if $p_\ell(n) \xrightarrow[\ell\to\infty]{} 0$ for all $n$. In turn, $p_\ell(n) \in 2^{\mathbb{N}}$, so saying that it converges to 0 means that $p_\ell(n)(m) \xrightarrow[\ell\to\infty]{} 0$ for all $m$. This last limit is a limit of numbers, where the only possible numbers are 0 and 1. This means that "converges to zero" and "is eventually equal to zero" are the same in this context.

**7.6.4. Proposition.**

*The space $2^{\mathbb{N}}$ is homeomorphic to $(2^{\mathbb{N}})^{\mathbb{N}}$. As a consequence, $\mathcal{C}$ is homeomorphic to $\mathcal{C}^{\mathbb{N}}$.*

*Proof.* Let $\mathbb{N} = \bigcup_k N_k$ be a partition in infinitely many infinite disjoint sets, $N_k = \{n_{k,r}\}_r$. Define $\eta : (2^{\mathbb{N}})^{\mathbb{N}} \to 2^{\mathbb{N}}$ by

$$\eta(\{g_m\})(n_{k,r}) = g_k(r).$$

This is a bijection, and it is also continuous (Exercise 7.6.1). Since the domain is compact, it is a homeomorphism.

By Proposition 7.6.2, there is a homeomorphism $\lambda : \mathcal{C} \to 2^{\mathbb{N}}$. Then, using Exercise 7.6.2,

$$\gamma : \mathcal{C} \xrightarrow{\lambda} 2^{\mathbb{N}} \xrightarrow{\eta} (2^{\mathbb{N}})^{\mathbb{N}} \xrightarrow{(\lambda^{-1})^{\mathbb{N}}} \mathcal{C}^{\mathbb{N}}$$

is a homeomorphism. □

Below is the main result of this section. One should appreciate how unexpected it is, since $T$ is *any* compact metrizable topological space! This should be seen as an example of the power of the ideas exposed in this chapter.

**7.6.5. Theorem.**

*Let $T$ be a compact and metrizable topological space. Then there is a continuous surjection $\alpha : \mathcal{C} \to T$.*

*Proof.* By Proposition 7.4.24 we have that $C(T)$ is separable. Since $B_1^{C(T)}(0)$ is separable (Proposition 1.8.5), there exists a countable dense subset $\{f_n\}$. Define $\nu : T \to \overline{\mathbb{D}}^{\mathbb{N}}$ by $\nu(t) = \{f_n(t)\}_n$. If $\nu(t) = \nu(s)$, then $f_n(t) = f_n(s)$ for all $n$; as any $f \in B_1^{C(T)}(0)$ is a limit of some of the $f_n$, we get that $f(t) = f(s)$ for all $f$ with $\|f\| \leq 1$, and by scaling we get that $f(t) = f(s)$ for all $f \in C(T)$. As $T$ is compact Hausdorff, Urysohn's Lemma (Theorem 2.6.5) implies that $s = t$, so $\nu$ is injective. As for continuity, if $t_j \to t$, then $f_n(t_j) \to f_n(t)$ for all $n$; as convergence in $\overline{\mathbb{D}}^{\mathbb{N}}$ is pointwise, this means that $\nu(t_j) \to \nu(t)$ and so $\nu$ is continuous. By Exercise 1.8.37, $\nu(T)$ is compact and by Exercise 1.8.38 it is a homeomorphism onto its image.

Since we can continuously embed $\overline{\mathbb{D}} \to [-1,1]^2 \to [0,1]^2$ as a homeomorphism onto its image (explicitly: $a+ib \longmapsto \left(\frac{a+1}{2}, \frac{b+1}{2}\right)$), we obtain a continuous injection $\tilde{\nu} : T \to [0,1]^{\mathbb{N}}$, homeomorphic onto its image (Exercise 7.6.3).

We have, from Proposition 7.6.4, Proposition 2.2.2, and Exercise 7.6.2,

$$\mathcal{C} \xrightarrow{\eta} \mathcal{C}^{\mathbb{N}} \xrightarrow{\beta^{\mathbb{N}}} [0,1]^{\mathbb{N}},$$

where the first map is a homeomorphism and the second one a continuous surjection. Then $(\beta^{\mathbb{N}} \circ \eta)^{-1}(\tilde{\nu}(T))$ is a closed subspace of $\mathcal{C}$. By Proposition 7.6.3, there exists a continuous retract $\gamma : \mathcal{C} \to (\beta^{\mathbb{N}} \circ \eta)^{-1}(\tilde{\nu}(T))$. Then

$$\alpha : \mathcal{C} \xrightarrow{\gamma} (\beta^{\mathbb{N}} \circ \eta)^{-1}(\tilde{\nu}(T)) \xrightarrow{\beta^{\mathbb{N}} \circ \eta} \tilde{\nu}(T) \xrightarrow{\tilde{\nu}^{-1}} T$$

is a continuous surjection. □

A remarkable consequence of Theorem 7.6.5 is

**7.6.6. Theorem. (Banach–Mazur)**

*The space $C[0,1]$ is universal as a separable Banach space, in the sense that every separable Banach space embeds isometrically into it.*

*Proof.* By Proposition 7.2.24, it is enough to show that $C(K)$ embeds isometrically into $C[0,1]$ when $K$ is a compact Hausdorff metric space.

Let $\{x_n\} \subset K$ be a countable dense subset. Let $d$ be a metric for $K$ such that $0 \leq d \leq 1$ (this can always be achieved by taking a metric $d'$ and using the equivalent metric $d = \frac{d'}{d'+1}$). Define $\rho : K \to [0,1]^{\mathbb{N}}$ by $\rho(x) = (d(x, x_n))_n$. Since we use the product topology on $[0,1]^{\mathbb{N}}$, which is pointwise convergence, $\rho$ is continuous since each component $x \longmapsto d(x, x_n)$ is continuous. If $x \neq y$, then there exists $n$ such that $d(x, x_n) < d(y, x_n)$ (by taking $x_n$ very close to $x$); thus $\rho(x) \neq \rho(y)$ and so $\rho$ is injective. As $K$ is compact and $[0,1]^{\mathbb{N}}$ is Hausdorff, it follows by Exercise 1.8.38 that $\rho$ is a homeomorphism onto its image. By Theorem 7.6.5 there exists a continuous surjection $\alpha : \mathcal{C} \to [0,1]^{\mathbb{N}}$. Let $R_0 = \alpha^{-1}(\rho(K)) \subset \mathcal{C}$. Then $R_0$ is compact and $\rho^{-1} \circ \alpha : R_0 \to K$ is a continuous surjection. By Proposition 7.4.9 there is an isometric homomorphism $\pi : C(K) \to C(R_0)$.

Since $R_0$ is closed, its complement $[0,1]\setminus R_0$ is a countable union of open intervals (Proposition 1.8.1), $R_0 = \bigcup_n (a_n, b_n)$. We can use this to define $T : C(R_0) \to C[0,1]$ by taking $g \in C(R_0)$ and joining, in the graph of $g$, the endpoints $(b_m, g(b_n))$ and $(a_{n+1}, g(a_{n+1}))$ with a line; this gives us a unique function $Tg \in C[0,1]$, and it makes $T$ linear. Clearly $\|Tg\| = \|g\|$, so $T$ is an isometric linear map. Then $T\circ\pi : C(K) \to C[0,1]$ is linear and isometric. □

One should marvel at the Banach–Mazur Theorem. It shows in an overwhelming way how vast infinity is. The space $C[0,1]$ looks simple enough, after all every continuous function there is a uniform limit of polynomials, as we showed in Theorem 7.4.20. But one can embed in it, isometrically, the space $C_0(\mathbb{R}^n)$ for any $n$. Or the space $\mathcal{K}(\mathcal{H})$ of the compact operators in a Hilbert space. Or any of the incredibly vast family of separable Banach spaces. And any normed space can be completed to a Banach space, so any separable normed space embeds isometrically into $C[0,1]$. The situation is radically different, of course, if instead of looking at $C[0,1]$ as a Banach space we look at it as a Banach algebra, or a C$^*$-algebra; no universality in those cases.

We finish the section with a result that will be useful down the road. For a function to be surjective it is necessary and sufficient that it has a right-inverse (Exercise 1.1.6). But such inverse need not be as good as the original function. Concretely, surjective continuous functions exist, between compact topological spaces, without a continuous right inverse (Exercise 2.2.13). But the next best thing holds:

***7.6.7. Proposition.***

*Let $S, T$ be compact metric spaces and $f : S \to T$ continuous and surjective. Then there exists $g : T \to S$, Borel, such that $f\circ g = \mathrm{id}_T$.*

*Proof.* By Theorem 7.6.5 there exists $\alpha : \mathcal{C} \to S$ continuous and surjective. Then $f\circ\alpha : \mathcal{C} \to T$ is continuous and surjective. Define $h : T \to \mathcal{C}$ by

$$h(t) = \inf (f\circ\alpha)^{-1}(\{t\}).$$

The function $h$ is Borel. Indeed, we will show that $h^{-1}(-\infty, a]$ is closed (hence Borel) for all $a \in \mathbb{R}$. Suppose that $\{t_n\} \subset h^{-1}(-\infty, a]$ is Cauchy. Since $T$ is compact it is complete (Proposition 1.8.21), so there exists $t \in T$ with $t_n \to t$. We have $h(t_n) \le a$; by definition of infimum there exists $c_n \in \mathcal{C}$ with $(f\circ\alpha)(c_n) = t_n$ and $c_n \le h(t_n) + \frac{1}{n} \le a + \frac{1}{n}$. Because $\mathcal{C}$ is a compact subset of $\mathbb{R}$ it is complete, and so there exists a convergent subsequence $\{c_{n_k}\}$ with $c_{n_k} \to c \in \mathcal{C}$. Then $c \le a$ and

$$(f\circ\alpha)(c) = \lim_k (f\circ\alpha)(c_{n_k}) = \lim_k t_{n_k} = t.$$

Hence $c \in (f\circ\alpha)^{-1}(t)$ with $c \leq a$, which implies that $h(t) \leq a$. So $h^{-1}(-\infty, a]$ is closed and as the sets $(-\infty, a]$ generated the Borel topology on $\mathbb{R}$, the function $h$ is Borel.

We also have, again by definition of infimum, that $h(t) = \lim_n d_n$, with $(f \circ \alpha)(d_n) = t$ for all $n$. Then, again by the continuity of $f \circ \alpha$,

$$(f \circ \alpha) \circ h(t) = \lim_n (f \circ \alpha)(d_n) = t.$$

We can now take $g = \alpha \circ h$. Then $g : T \to S$ is Borel (Proposition 2.4.6) and $f \circ g(t) = t$ for all $t$. $\square$

### 7.6.1. Exercises.

(7.6.1) Show that $\eta$, as in the proof of Proposition 7.6.4, is a continuous bijection.

(7.6.2) Let $X, Y$ be topological spaces and $\alpha : X \to Y$ a homeomorphism. Define $\alpha^{\mathbb{N}} : X^{\mathbb{N}} \to Y^{\mathbb{N}}$ by $\alpha^{\mathbb{N}}(\{x_n\}) = \{\alpha(x_n)\}$. Show that $\alpha^{\mathbb{N}}$ is a homeomorphism.

(7.6.3) In the proof of Theorem 7.6.5, it is claimed that using $\nu : T \to \overline{\mathbb{D}}^{\mathbb{N}}$ and the homeomorphic embedding of $\overline{\mathbb{D}}$ into $[0,1]^2$ one gets a continuous injective $\tilde{\nu} : T \to [0,1]^{\mathbb{N}}$, homeomorphic onto its image. Write the details to justify these assertions.

CHAPTER

# 8

# Fourier Series and Hilbert Function Spaces

The topics in this chapter are not properly considered Functional Analysis. But they are connected, both for theoretical reasons and as a source of examples. So while the idea is to provide technical detail, we will also restrict ourselves to just the required level of generality to achieve some results, to simplify and shorten the exposition a bit. In particular, we will only discuss Fourier series on the interval $[-\pi, \pi]$ (or in the unit circle, depending on point of view). All results translate to Fourier series on other intervals. It is worth mentioning that Fourier analysis can be done in much more generality, in the context of locally compact abelian groups, but here we will stick to the classical setting.

## 8.1. Fourier Series

The origin of the study of Fourier series is the idea, inspired by Physics, that functions should be expressible as superposition of waves; that is, as (limits of) linear combinations of sines and cosines. Classically, a Fourier series on the interval $[-\pi,\pi]$ looks like

$$f(x)=\sum_{n=0}^{\infty}a_n\cos nx+\sum_{n=1}^{\infty}b_n\sin nx.$$

It is common to consider the above function as real, but we have been using complex coefficients already. Then one notices that

$$\operatorname{span}\{\cos nx:\ n\in\mathbb{N}\}\cup\{\sin nx:\ n\in\mathbb{N}\}=\operatorname{span}\{e^{inx}:\ n\in\mathbb{Z}\},\tag{8.1}$$

and that in both cases (after normalizing properly) what we are dealing with are orthogonal bases of $L^2[-\pi,\pi]$.

By a slight abuse of language, we will consider any function $f:[-\pi,\pi]\to\mathbb{C}$ to extend periodically to the rest of $\mathbb{R}$. In particular, when $f$ is defined pointwise (as opposed to $f\in L^1[-\pi,\pi]$, say) we assume that $f(-\pi)=f(\pi)$.

### 8.1.1. Definition.

*Given $f\in L^1[-\pi,\pi]$, its **Fourier coefficients** are the numbers*

$$\hat f(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\,e^{-int}\,dt,\qquad n\in\mathbb{Z}.$$

*The **Fourier series** of $f$ is the formal series*

$$S_f(t)=\sum_{n\in\mathbb{Z}}\hat f(n)\,e^{int}.\tag{8.2}$$

*Given $n\in\mathbb{N}$ we also consider the corresponding partial sum of $S_f$, that is*

$$S_{f,n}(t)=\sum_{k=1}^{n}\hat f(k)\,e^{ikt}.$$

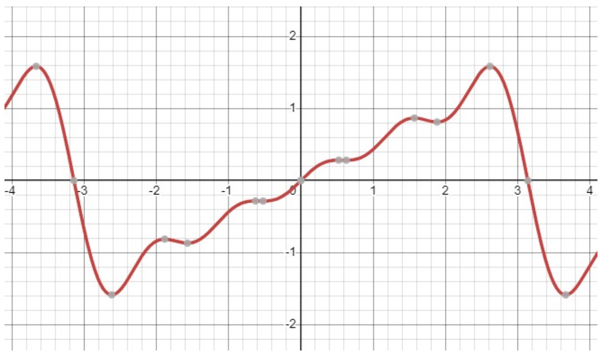
If we consider $f\in L^2[-\pi,\pi]\subset L^1[-\pi,\pi]$, the Fourier coefficients of $f$ are—up to a fixed scalar multiple—simply the coefficients of $f$ with respect to the orthonormal basis $\{e^{int}\}_{n\in\mathbb{Z}}$ in $L^2[-\pi,\pi]$ with the normalized measure $\frac{1}{2\pi}\,m$, and the map $f\longmapsto\hat f$ is an explicit Hilbert space isomorphism (that is, a **unitary**) between $L^2[-\pi,\pi]$ and $\ell^2(\mathbb{Z})$.

While it still remains to be seen to what extent the Fourier series of a function actually represents the function, here are some basic examples. We will later prove that in all these examples the Fourier series converges to the function at most points.

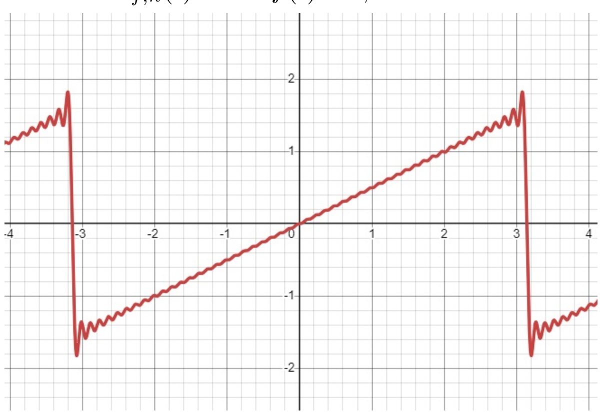
Consider first the function $f(t) = t$. We are restricting ourselves to $t \in [-\pi, \pi]$. As trigonometric functions are sinusoidal and periodic, one would expect that it should be kind of hard to approximate $f$ with trigonometric polynomials. We have (Exercise 8.1.1) that

$$S_{f,n}(t) = \sum_{k=1}^{n} \frac{(-1)^{k+1} \sin kt}{k}.$$

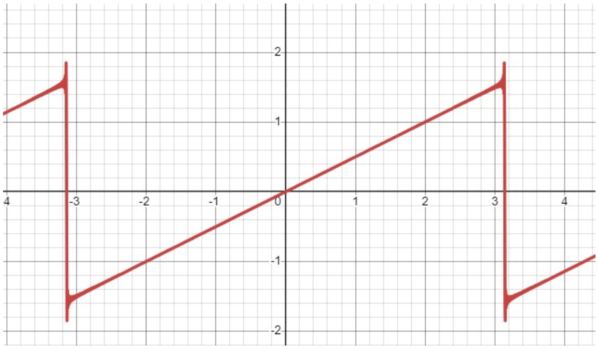
Below we graph $S_{f,n}(t)$ for $n = 5, 50, 500$:




$S_{f,n}(t)$ when $f(t) = t$, $\qquad n = 5$




$S_{f,n}(t)$ when $f(t) = t$, $\qquad n = 50$

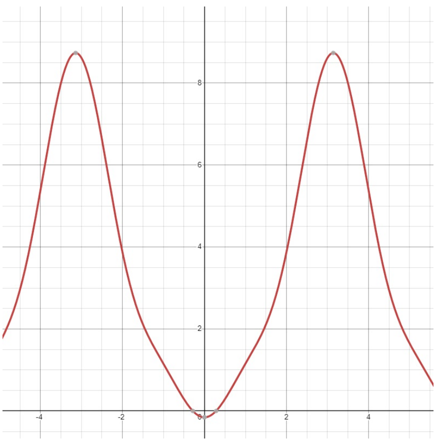



$S_{f,n}(t)$ when $f(t) = t$, $\qquad n = 500$

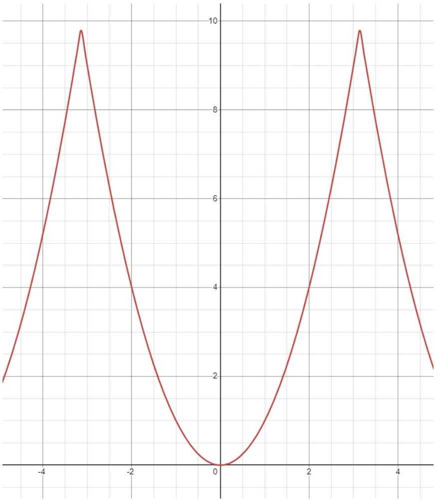
It should not be a surprise that trigonometric polynomials have issues with approximating sharp corners. With that in mind, approximating $f(t) = t^2$ should be easier, and in fact it is. We have

$$S_{f,n}(t) = \frac{\pi^2}{3} + \sum_{k=1}^{n} \frac{4(-1)^k \cos kt}{k^2}$$

(Exercise 8.1.2). We graph below $S_{f,n}(t)$ for $n = 3$ and $n = 50$:



$S_{f,n}(t)$ when $f(t) = t^2$, $\qquad n = 3$



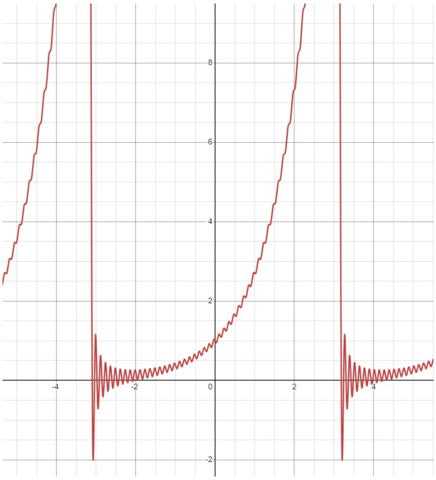
$S_{f,n}(t)$ when $f(t) = t^2$, $\qquad n = 50$

Of course, the approximation only works up on $[-\pi, \pi]$, since outside of that our trigonometric polynomials extend periodically. And as a last example let us consider something a bit harder to approximate with trigonometric polynomials. If $f(t) = e^t$, then

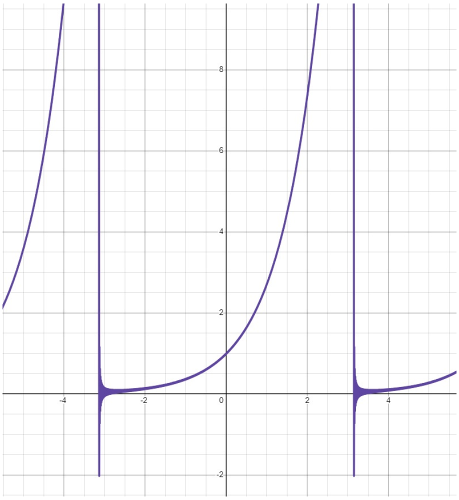
$$S_{f,n}(t) = \frac{\sinh \pi}{2\pi} + \sum_{k=1}^{n} \frac{2(-1)^k \sinh \pi \left( \cos kt - k \sin kt \right)}{\pi(1 + k^2)}$$

([Exercise 8.1.3](#)). We graph below $S_{f,n}(t)$ for $n = 5$, $n = 50$, and $n = 500$:


$S_{f,n}(t)$ when $f(t) = e^t$, $\quad n = 5$




$S_{f,n}(t)$ when $f(t) = e^t$, $\quad n = 50$




$S_{f,n}(t)$ when $f(t) = e^t$, $\quad n = 500$

Our first goal is to establish that the trigonometric basis is indeed enough to somehow describe all functions in $L^2[-\pi, \pi]$, and eventually in $L^1[-\pi, \pi]$. It remains to be seen what we mean by "describe".

**8.1.2. Lemma.**

*Let $q \in \mathbb{C}[x, y]$. Then $q(\cos t, \sin t) \in \operatorname{span}\{e^{int} : n \in \mathbb{Z}\}$.*

*Proof.* It is enough to show it for a monomial. We have

$$\begin{aligned}\cos^m t \sin^n t &= \left(\frac{e^{it}+e^{-it}}{2}\right)^m \left(\frac{e^{it}-e^{-it}}{2i}\right)^n \\ &= 2^{-m-n}\,(-i)^n \sum_{k=0}^{m} \binom{m}{k}\, e^{i(2k-m)t} \sum_{j=0}^{n} \binom{n}{j}\, e^{i(2j-n)t} \\ &= 2^{-m-n}\,(-i)^n \sum_{k=0}^{m}\sum_{j=0}^{n} \binom{m}{k}\binom{n}{j}\, e^{i(2k+2j-m-n)t} \\ &\in \operatorname{span}\{e^{int} : n \in \mathbb{Z}\}.\end{aligned}$$

□

**8.1.3. Proposition.**

*The trigonometric system $\{e^{int}\}_{n\in\mathbb{Z}}$ is complete (that is, it is maximal orthogonal). More generally, if $f \in L^1[-\pi,\pi]$ and $\hat{f}(n) = 0$ for all $n$, then $f = 0$. If $f \in L^2[-\pi,\pi]$, then*

$$\|f\|_2 = \sqrt{2\pi}\left(\sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2\right)^{1/2}. \tag{8.3}$$

*Proof.* Suppose first that $f \in C[-\pi,\pi]$, real valued, and $\hat{f}(n) = 0$ for all $n$. We have

$$\int_{-\pi}^{\pi} f(t)\,\cos t\,dt = \frac{1}{2}\int_{-\pi}^{\pi} f(t)\,e^{it}\,dt + \frac{1}{2}\int_{-\pi}^{\pi} f(t)\,e^{-it}\,dt = 0+0 = 0.$$

Similarly we get that $\int_{-\pi}^{\pi} f(t)\,\sin t\,dt = 0$. By linearity and Lemma 8.1.2,

$$\int_{-\pi}^{\pi} f(t)\,q(\cos t, \sin t)\,dt = 0 \tag{8.4}$$

for all $q \in \mathbb{C}[x,y]$. If $f \neq 0$, then $|f|$ attains a maximum $M$ on $[-\pi,\pi]$, say at $t_0$. Choose $\delta > 0$ such that $\delta < \frac{\pi}{2}$ and $|f(t)| \geq \frac{M}{2}$ for $|t-t_0| < \delta$, and let

$$p(t) = 1 + \cos(t-t_0) - \cos\delta.$$

As $\cos(t-t_0) = \cos t_0\,\cos t - \sin t_0\,\sin t$, we get by the computations above that

$$\int_{-\pi}^{\pi} f(t)\,p(t)^n\,dt = 0, \qquad n \in \mathbb{N}.$$

When $|t-t_0| < \delta/2$ we have $\cos(t-t_0) \geq \cos\frac{\delta}{2} - \cos\delta > 0$, and so $p(t) \geq 1+s$, where $s = \cos\frac{\delta}{2} - \cos\delta$. Then

$$\int_{-\pi}^{\pi} f(t)\,p(t)^n\,dt = \int_{|t-t_0|\leq\delta} f(t)\,p(t)^n\,dt + \int_{|t-t_0|>\delta} f(t)\,p(t)^n\,dt$$

$$\geq \int_{|t-t_0|\leq\frac{\delta}{2}} f(t)\,p(t)^n\,dt + \int_{|t-t_0|>\delta} f(t)\,p(t)^n\,dt$$
$$\geq (1+s)^n\,\frac{\delta M}{4} - 2\pi M,$$

where we are using that $f(t) \geq -|f(t)| \geq -M$, and $|p(t)| \leq 1$ if $|t-t_0| \geq \delta$. The last term above is unbounded on $n$, contradicting (8.4). If follows that $f = 0$.

If now $f \in C(T)$,

$$\begin{aligned}\int_{-\pi}^{\pi} \operatorname{Re} f(t)\,e^{int}\,dt &= \frac{1}{2}\int_{-\pi}^{\pi} \left[f(t) + \overline{f(t)}\right] e^{int}\,dt \\ &= \frac{1}{2}\int_{-\pi}^{\pi} f(t)\,e^{int}\,dt + \frac{1}{2}\overline{\int_{-\pi}^{\pi} f(t)\,e^{-int}\,dt}\end{aligned}$$

so $\hat f(n) = 0$ for all $n$ forces $\widehat{\operatorname{Re} f}(n) = 0$ for all $n$, and then $\operatorname{Re} f = 0$ by the first paragraph. The same idea works for the imaginary part, so we get that $f = 0$.

When $f \in L^1[-\pi,\pi]$, let $g(x) = \int_0^x f$. Then $g$ is continuous by Exercise 2.8.10. We have, by Fubini (Theorem 2.7.16), since $f$ is integrable and the intervals are bounded, for $n \neq 0$

$$\begin{aligned}2\pi\,\hat g(n) &= \int_{-\pi}^{\pi}\int_0^x f(t)\,e^{-inx}\,dt\,dx = \int_{-\pi}^{\pi}\int_t^{\pi} f(t)\,e^{-inx}\,dx\,dt \\ &= -\frac{1}{in}\int_{-\pi}^{\pi} f(t)\left[(-1)^n - e^{-int}\right]dt = \frac{(-1)^{n+1}}{in}\,\hat f(0) + i\,\frac{1}{n}\,\hat f(n) = 0.\end{aligned}$$

Then the function $h = g - \hat g(0)$ is continuous with $\hat h(n) = 0$ for all $n$. By the first part of the proof $h = 0$, so $g$ is constant. And then $f = g' = 0$ a.e. by the Fundamental Theorem of Calculus (Corollary 2.11.14). We have thus shown that for $f \in L^1[-\pi,\pi]$, $f = 0$ if and only if $\hat f = 0$. Since $L^2[-\pi\pi] \subset L^1[-\pi,\pi]$, we have also shown that the orthogonal system $\{e^{int}\}_{n\in\mathbb{Z}}$ is complete. Thus $\{(2\pi)^{-1/2}e^{int}\}_{n\in\mathbb{Z}}$ is orthonormal, and so by Parseval's equality (Corollary 4.3.23)

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\langle f,(2\pi)^{-1/2}e^{int}\rangle|^2 = \frac{1}{2\pi}\sum_{n\in\mathbb{Z}} |2\pi\,\hat f(n)|^2 = 2\pi\sum_{n\in\mathbb{Z}} |\hat f(n)|^2. \qquad \square$$

One can work in more generality and use an arbitrary orthonormal basis of $L^2[-\pi,\pi]$, but the trigonometric monomials are the canonical choice and we will stick with that. One advantage of this particular system is that its functions are bounded, which allow us to define the Fourier coefficients for $f \in L^1[-\pi,\pi]$.

With a Fourier series we are representing a function $f$ by a countable set of numbers $\{\hat f(n)\}$. When the Fourier series $S_f$ approximates $f$ in some sense, we can the describe this approximation with finitely many numbers $\hat f(1),\ldots,\hat f(m)$. This is the basic idea behind digital sound storage/transmission.

The reason Fourier series are not just an exercise in Chapter 4 is that, due to the fact that we are dealing with functions, we can ask ourselves whether we have other convergence than the 2-norm in (8.2); we can also consider the Fourier series of functions not in $L^2$, so it makes sense to wonder in what cases (8.2) actually represents $f$, and how.

Concretely, the previous paragraph establishes that $f$ is determined by its Fourier coefficients, as long as $f\in L^2[-\pi,\pi]$ and—crucially—if we don't mind writing $f$ as a $\|\cdot\|_2$-limit of partial sums. But what if $f$ is in $L^1$ but not in $L^2$? And, can we get a better convergence that the 2-norm, in order to actually calculate values of $f$? We may consider the series in (8.2) and ask if the series converges pointwise, if it converges pointwise almost everywhere, if it converges uniformly, or in some other norms. And there is a further question: if the Fourier series converges in some sense, does it converge to $f$?

**8.1.1. Pointwise Convergence.** We begin with a best-case scenario.

***8.1.4. Proposition.***

*Let $f\in C[-\pi,\pi]$ be continuously differentiable and such that $f(-\pi)=f(\pi)$. Then $S_f(t)$ converges uniformly, and it is equal to $f$ at every point. Also $S_{f'}(t)=S_f(t)'$.*

*Proof.* Let us calculate the Fourier coefficients of $f'$. Integrating by parts,

$$\widehat{f'}(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}f'(t)\,e^{-int}\,dt=\frac{f(t)\,e^{-int}}{2\pi}\Big|_{-\pi}^{\pi}+\frac{in}{2\pi}\int_{-\pi}^{\pi}f(t)\,e^{-int}\,dt=in\,\hat f(n).$$

As $f'\in L^2[-\pi,\pi]$, Parseval's Equality gives us $\sum_n n^2|\hat f(n)|^2=\sum_n|\widehat{f'}(n)|^2<\infty$. Then, using Cauchy–Schwarz,

$$\sum_{n\neq0}|\hat f(n)|=\sum_{n\neq0}|n\hat f(n)|\,\frac1n\leq\Big(\sum_{n\neq0}n^2|\hat f(n)|^2\Big)^{1/2}\Big(\sum_{n\neq0}\frac{1}{n^2}\Big)^{1/2}<\infty.$$

So the Fourier series $\sum_n\hat f(n)\,e^{int}$ converges absolutely and uniformly to a continuous function. This function has the same Fourier coefficients as $f$, so it agrees with $f$ almost everywhere; with both function being continuous, they are equal. ☐

The condition $f(-\pi)=f(\pi)$ looks artificial, but it is satisfied by any $S_f(t)$, so only such functions can be equal pointwise to their Fourier series. Let us now get some information about the coefficients $\hat f(n)$.

**8.1.5. Lemma. (Riemann–Lebesgue)**

*Let $f \in L^1[-\pi,\pi]$. Then* $\lim_{n\to\pm\infty} \hat{f}(n) = 0$.

*Proof.* Assume first that $f$ is continuously differentiable, and that $f(-\pi) = f(\pi)$. By Proposition 8.1.4, $|\hat{f}(n)| \to 0$. Now for arbitrary integrable $f$, fix $\varepsilon > 0$. By Proposition 2.8.24 there exists $g$, continuously differentiable, with $g(-\pi) = g(\pi)$ (we can achieve this condition by adding to $g$ a continuous function with very small support) and $\|f-g\|_1 < \varepsilon$. Then

$$\begin{aligned}|\hat{f}(n)| &\leq |\hat{f}(n) - \hat{g}(n)| + |\hat{g}(n)| \\ &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} [f(t) - g(t)]\, e^{-int}\, dt\right| + |\hat{g}(n)| \leq \varepsilon + |\hat{g}(n)|.\end{aligned}$$

So $\limsup_{n\to\pm\infty} |\hat{f}(n)| \leq \varepsilon$. By the Limsup Routine, the limit exists and is zero. □

**8.1.6. Lemma. (Taibleson)**

*Let $f : [-\pi,\pi] \to \mathbb{C}$ be of bounded variation, and let $T$ denote its total variation. Then*

$$|\hat{f}(n)| \leq \frac{2\pi T}{|n|}, \qquad n \in \mathbb{N}.$$

*Proof.* As we can work with the real and imaginary parts of $f$, we may assume without loss of generality that $f$ is real-valued. Let $a_j = \frac{2(j-|n|)\pi}{|n|}$ be a partition of $[-\pi,\pi]$ with equal lengths. Define

$$g = \sum_{j=1}^{n} f(a_j)\, 1_{(a_{j-1},a_j)}.$$

Then, since $e^{-ina_j} = 1$ for all $j$,

$$\begin{aligned}\int_{-\pi}^{\pi} g(t)\, e^{-int}\, dt &= \sum_{j=1}^{n} f(a_j) \int_{a_{j-1}} a_j e^{-int}\, dt \\ &= \sum_{j=1}^{n} f(a_j)(e^{-ina_j} - e^{-ina_{j-1}}) = 0.\end{aligned}$$

Thus

$$|\hat{f}(n)| = \left|\int_{-\pi}^{\pi} f(t)\, e^{-int}\, dt\right| = \left|\int_{-\pi}^{\pi} \big[f(t) - g(t)\big]\, e^{-int}\, dt\right|$$

$$= \left|\sum_{j=1}^{|n|}\int_{a_{j-1}}^{a_j}\big[f(t)-f(a_j)\big]\,e^{-int}\,dt\right|$$
$$\le \sum_{j=1}^{|n|}\int_{a_{j-1}}^{a_j}\big|f(t)-f(a_j)\big|\,dt$$
$$\le \sum_{j=1}^{|n|}\sup\big\{\big|f(t)-f(a_j)\big|:\ t\in[a_{j-1},a_j]\big\}\,|a_j-a_{j-1}|$$
$$= \frac{2\pi}{|n|}\sum_{j=1}^{|n|}\sup\big\{\big|f(t)-f(a_j)\big|:\ t\in[a_{j-1},a_j]\big\}\le\frac{2\pi T}{|n|}. \qquad \square$$

*8.1.7. Example.* Here is a byproduct of the study of Fourier Series. Consider $f(t)=t$ on $L^2[-\pi,\pi]$. Its Fourier coefficients are

$$\hat f(n)=\frac{1}{2\pi}\int_{-\pi}^{\pi}t\,e^{int}\,dt=\frac{t\,e^{int}}{2in\pi}\Big|_{-\pi}^{\pi}-\frac{1}{2in\pi}\int_{-\pi}^{\pi}e^{int}\,dt=\frac{(-1)^{n+1}}{in}, \qquad n\ne 0$$

and $\hat f(0)=0$. Then Proposition 8.1.3 gives us

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac12\sum_{n\ne0}\frac{1}{|in|^2}=\frac12\,\frac{1}{2\pi}\int_{-\pi}^{\pi}t^2\,dt=\frac{2\pi^3}{12\pi}=\frac{\pi^2}{6},$$

with a bit less work than in Example 3.10.6.

For real valued functions one may want to not have imaginary numbers hanging around, so it is common to write the Fourier series in terms of $\cos nt$ and $\sin nt$ (this is doable because of (8.1)). If we have $S_f(t)=\sum_{n\in\mathbb Z}c_ne^{int}$ and $f$ is real we have $c_{-n}=\overline{c_n}$, so if $c_n=a_n+ib_n$

$$S_f(t)=\sum_{n\in\mathbb Z}c_ne^{int}=c_0+\sum_{n\in\mathbb N}2\mathrm{Re}\,c_ne^{int}=c_0+\sum_{n\in\mathbb N}2a_n\cos nt-2b_n\sin nt.$$

The convention is to scale the coefficients so that

$$S_f(t)=\frac12\,a_0+\sum_{n\in\mathbb N}a_n\cos nt+b_n\sin nt.$$

Let us now manipulate the partial sums of $S_f(t)$. We will denote by $S_{f,n}(t)$ the partial sum $S_{f,n}(t)=\sum_{|k|\le n}\hat f(k)\,e^{ikt}$. We have

$$S_{f,n}(t)=\sum_{k=-n}^{n}c_ke^{ikt}=\sum_{k=-n}^{n}\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\,e^{ik(t-x)}\,dx$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}\operatorname{Re} e^{ik(t-x)}\right)dx$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos k(x-t)\right)dx$$

$$= \int_{-\pi}^{\pi} f(t)\,D_n(x-t)\,dx = \frac{1}{\pi}\,(f*D_n)(t),$$

where

$$D_n(t) = \frac{1}{2}+\sum_{k=1}^{n}\cos kt$$

is the $n^{\text{th}}$ **Dirichlet Kernel**. We can write this rather explicitly by noting that, for $t \neq 0$,

$$D_n(t) = \frac{1}{2}+\sum_{k=1}^{n}\cos kt = \frac{1}{2}+\operatorname{Re}\sum_{k=1}^{n}e^{ikt} = \frac{1}{2}+\operatorname{Re}\frac{e^{it}-e^{i(n+1)t}}{1-e^{it}}$$

$$= \frac{1}{2}+\operatorname{Re}\frac{e^{it}-e^{i(n+1)t}}{1-e^{it}}\,\frac{1-e^{-it}}{1-e^{-it}} = \frac{1}{2}+\operatorname{Re}\frac{e^{it}-1-e^{i(n+1)t}+e^{int}}{2(1-\cos t)}$$

$$= \frac{1}{2}+\frac{-1+\cos t-\cos(n+1)t+\cos nt}{2(1-\cos t)} = \frac{-\cos(n+1)t+\cos nt}{2(1-\cos t)}$$

$$= \frac{-\cos\left[(n+\frac{1}{2})t+\frac{t}{2}\right]+\cos[(n+\frac{1}{2})t-\frac{t}{2}]}{2(1-\cos 2\frac{t}{2})}$$

$$= \frac{-\cos(n+\frac{1}{2})t\cos\frac{t}{2}+\sin(n+\frac{1}{2})t\,\sin\frac{t}{2}+\cos(n+\frac{1}{2})t\,\cos\frac{t}{2}+\sin(n+\frac{1}{2})t\,\sin\frac{t}{2}}{4\sin^2\frac{t}{2}}$$

$$= \frac{\sin(n+\frac{1}{2})t}{2\sin\frac{t}{2}}$$

Using the function $f(t)=1$ we have that $S_{f,n}(t)=1$ for all $n$, and so we find that $\displaystyle\int_{-\pi}^{\pi} D_n(t)\,dt = \pi$ for all $n$.

### 8.1.2. Cesàro Convergence.

**8.1.8. Definition.**

*Let $\{a_n\}\subset\mathbb{C}$ be a sequence. We say that $\{a_n\}$ **converges Cesàro** to $a$ if*

$$\lim_{n\to\infty}\frac{a_1+\cdots+a_n}{n} = a.$$

It is well-known that Cesàro convergence is strictly weaker than usual convergence, and strictly stronger than Abel convergence when applied to partial sums of series.

**8.1.9. Proposition.**

*Let $\{a_n\} \subset \mathbb{C}$ be a sequence such that $\sum_n a_n$ converges Cesàro to $a$. If there exists $c > 0$ such that $|a_n| \leq \frac{c}{n}$ for all $n$, then $\sum_n a_n = a$.*

*Proof.* Fix $\varepsilon > 0$ and put $k_n = \lfloor \varepsilon n \rfloor$. Let $s_n = \sum_{k=1}^n a_k$, $c_n = \frac{1}{n} \sum_{k=1}^n s_n$. The hypothesis is that $c_n \to a$. Let

$$d_n = \frac{s_{n+1} + \cdots + s_{n+k_n}}{k_n}.$$

Then

$$\begin{aligned} d_n &= \frac{s_1 + \cdots + s_{n+k_n} - (s_1 + \cdots + s_n)}{k_n} = \frac{n + k_n}{k_n} c_{n+k_n} - \frac{n}{k_n} c_n \\ &= \frac{n}{k_n} (c_{n+k_n} - c_n) + c_{n+k_n}. \end{aligned}$$

As $\frac{n}{k_n} < \frac{2}{\varepsilon}$, we deduce that $d_n \to a$. We also have, for $n$ big enough so that $k_n > 0$,

$$\begin{aligned} |d_n - s_n| &= \left| \frac{(s_{n+1} - s_n) + \cdots + (s_{n+k_n} - s_n)}{k_n} \right| = \frac{1}{k_n} \left| \sum_{j=1}^{k_n} (k_n - j + 1) a_{n+k_n} \right| \\ &\leq \sum_{j=1}^{k_n} \left(1 - \frac{j-1}{k_n}\right) |a_{n+j}| \leq \sum_{j=1}^{k_n} |a_{n+k_n}| \\ &\leq \sum_{j=1}^{k_n} \frac{c}{n+j} \leq \frac{ck_n}{n+1} \leq c\,\varepsilon. \end{aligned}$$

It follows that $\limsup |d_n - s_n| \leq c\varepsilon$; By the Limsup Routine, $\lim_n |d_n - s_n| = 0$ and thus $\lim_n s_n = \lim_n d_n = a$. $\square$

**8.1.10. Definition.**

*Let $f : [-\pi, \pi] \to \mathbb{C}$ and let $S_{f,n}$ be the partial Fourier sums for $f$. We denote by*

$$C_{f,n}(t) = \frac{S_{f,0}(t) + \cdots + S_{f,n}(t)}{n+1}$$

*the partial Cesàro sums of $S_f$. We say that $S_f(t)$* ***converges Cesàro*** *to $g(t)$ if the sequence $\{C_{f,n}(t)\}$ converges to $g(t)$.*

In a similar way on how we obtained the Dirichlet kernel, we can write

$$C_{f,n}(t) = \frac{1}{n+1}\sum_{k=0}^{n} S_{f,k}(t) = \frac{1}{n+1}\sum_{k=0}^{n}\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, D_k(x-t)\, dt$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\, K_n(x-t)\, dt,$$

where

$$K_n(t) = \frac{1}{n+1}\sum_{k=0}^{n} D_k(t) = \frac{1}{n+1}\sum_{k=0}^{n}\frac{\sin(k+\frac{1}{2})t}{2\sin\frac{t}{2}}$$

is the **Fejér Kernel**. Using the equality $2\sin(k+\frac{1}{2})t\,\sin\frac{t}{2} = \cos kt - \cos(k+1)t$ the sum telescopes, and we get

$$K_n(t) = \frac{1-\cos(n+1)t}{4(n+1)\sin^2\frac{t}{2}} = \frac{2}{n+1}\left(\frac{\sin(n+1)\frac{t}{2}}{2\sin\frac{t}{2}}\right)^2.$$

In particular, $K_n(t) \geq 0$ for all $n, t$. As $K_n(t)$ is an average of the $D_k(t)$, we obtain immediately that

$$\int_{-\pi}^{\pi} K_n(t)\, dt = \pi. \tag{8.5}$$

Also, if $|t| \geq \delta$ then as $\sin\frac{t}{2}$ is monotone on $[-\pi, \pi]$,

$$K_n(t) = \frac{2}{n+1}\left(\frac{\sin(n+1)\frac{t}{2}}{2\sin\frac{t}{2}}\right)^2 \leq \frac{1}{2(n+1)\sin^2\frac{\delta}{2}}. \tag{8.6}$$

We also mention the trivial bound, obtained from $|D_k(t)| \leq n + \frac{1}{2}$,

$$K_n(t) \leq \frac{n\left(n+\frac{1}{2}\right)}{n+1} \leq n. \tag{8.7}$$

When it exists, we write $f(t-) = \lim_{x\to t^-} f(x)$, $f(t+) = \lim_{x\to t^+} f(x)$.

**8.1.11. Theorem. (Fejér)**

*Let $f \in L^1[-\pi, \pi]$. If $f$ is continuous at $t$ then $C_{f,n}(t) \to f(t)$, and the convergence is uniform on any closed interval in which $f$ is continuous. If $f$ has a jump discontinuity at $t$, then*

$$C_{f,n}(t) \xrightarrow[n\to\infty]{} \frac{f(t+)+f(t-)}{2}.$$

*Proof.* We first assume that $f$ is continuous on $[a, b] \subset [-\pi, \pi]$, where we allow $b = a$. Fix $\varepsilon > 0$ and $t \in [a, b]$. As $f$ is uniformly continuous on $[a, b]$, there exists $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$ when $|y - x| < \delta$. Then

$$\left|\frac{1}{\pi}\int_{|t|<\delta} [f(t-x) - f(t)]\, K_n(x)\, dx\right| \leq \frac{1}{\pi}\int_{|t|<\delta} \left|f(t-x) - f(t)\right| K_n(x)\, dx$$

$$\leq \frac{\varepsilon}{\pi}\int_{-\pi}^{\pi} K_n(x)\,dx = \varepsilon.$$

And, if $c = \max\{|f(x)| : \ x \in [a,b]\}$, $d = \max\{K_n(t) : \ \delta \leq |t| \leq \pi\}$, and using (8.5),

$$\begin{aligned}\left|\frac{1}{\pi}\int_{\pi\geq|t|\geq\delta}\big[f(t-x)-f(t)\big]\,K_n(x)\,dx\right| &\leq \frac{1}{\pi}\int_{\pi\geq|t|\geq\delta}\big[|f(t-x)|+c\big]K_n(x)\,dx\\ &\leq \frac{d}{\pi}\int_{\pi\geq|t|\geq\delta}\big[|f(t-x)|+c\big]\,dx\\ &\leq \frac{2cd}{(n+1)\sin^2\frac{\delta}{2}}\end{aligned}$$

Combining both estimates above and using (8.5),

$$\begin{aligned}|C_{f,n}(t)-f(t)| &= \left|\frac{1}{\pi}\int_{|t|<\delta}\big[f(t-x)-f(t)\big]\,K_n(x)\,dx\right|\\ &\quad + \left|\frac{1}{\pi}\int_{\pi\geq|t|\geq\delta}\big[f(t-x)-f(t)\big]\,K_n(x)\,dx\right|\\ &\leq \varepsilon + \frac{2cd}{(n+1)\sin^2\frac{\delta}{2}},\end{aligned}$$

and so $\limsup_n |C_{f,n}(t)-f(t)| \leq \varepsilon$. By the Limsup Routine, $\lim_n C_{f,n}(t) = f(t)$ uniformly on $[a,b]$, as the bounds do not depend on $t$, only on the continuity.

When there is a jump discontinuity at $t$, the proof works exactly as above with the difference that one assumes that $f(t) = \frac{1}{2}\big(f(t+)+f(t-)\big)$, and one splits the first integral into the integral on $[0,\delta]$ (where one uses the right limit) and the integral on $[-\delta,0]$, where one uses the left limit. □

***8.1.12. Corollary.***

*Let $f \in L^p[-\pi,\pi]$, $1 \leq p < \infty$. Then $\|C_{f,n}\|_p \leq \|f\|_p$, and $\|C_{f,n}-f\|_p \to 0$.*

*Proof.* We have, writing $1 = \frac{1}{p}+\frac{1}{q}$, and using (8.5) and Hölder,

$$\begin{aligned}\|C_{f,n}-f\|_p^p &= \int_{-\pi}^{\pi}\left|\frac{1}{\pi}\int_{-\pi}^{\pi}\big[f(x-t)-f(x)\big]\,K_n(t)\,dt\right|^p dx\\ &= \int_{-\pi}^{\pi}\left|\int_{-\pi}^{\pi}\frac{1}{\pi^{1/p}}\big[f(x-t)-f(x)\big]\,K_n(t)^{1/p}\,\frac{1}{\pi^{1/q}}K_n(t)^{1/q}\,dt\right|^p dx\end{aligned}$$

$$\leq \int_{-\pi}^{\pi} \left[\frac{1}{\pi}\int_{-\pi}^{\pi} \left|f(x-t)-f(x)\right|^p K_n(t)\,dt\right] \left[\frac{1}{\pi}\int_{-\pi}^{\pi} K_n(t)\right]^{p/q} dx$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left|f(x-t)-f(x)\right|^p K_n(t)\,dt\,dx$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left|f(x-t)-f(x)\right|^p K_n(t)\,dx\,dt$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi} \|f_t - f\|_p^p\, K_n(t)\,dt = C_{g,n}(0) \xrightarrow[n\to\infty]{} 0$$

by Theorem 8.1.11, since the function $g : t \longmapsto \|f_t - f\|_p$ is continuous at 0 by Lemma 2.8.19. We can exchange the integrals by Tonelli (Corollary 2.7.17), since the integrand is non-negative.

To show that $\|C_{f,n}\|_p \leq \|f\|_p$, we use the exact same estimates from above, applied only to $C_{n,f}$, to get

$$\|C_{f,n}\|_p^p \leq \frac{1}{\pi}\int_{-\pi}^{\pi} \|f\|_p^p\, K_n(t)\,dt = \|f\|_p^p. \qquad \square$$

We remark the easy fact that if $f$ is continuous then $\frac{1}{2}\big(f(t+)+f(t-)\big) = f(t)$. Now let us see that when $f$ is nice enough, the result in Theorem 8.1.11 can be improved to standard convergence.

***8.1.13. Corollary.***

*Let $f : [-\pi,\pi] \to \mathbb{C}$ be of bounded variation. Then*

*(i) for every $t \in [-\pi,\pi]$, $S_f(t)$ converges to $\frac{1}{2}\big(f(t+)+f(t-)\big)$;*

*(ii) if $f$ is continuous on $[a,b] \subset [-\pi,\pi]$, then $S_f$ converges uniformly to $f$ on $[a,b]$;*

*(iii) the partial sums of $S_f$ are bounded uniformly on both $t$ and $n$.*

*Proof.* Since $f$ is of bounded variation, its side limits exist (Exercise 2.11.7). We know that $f$ is integrable (Exercise 2.11.9). By Theorem 8.1.11, $S_f$ converges Cesàro (uniformly on closed intervals of continuity) to $\frac{1}{2}\big(f(t+) + f(t-)\big)$; and then Lemma 8.1.6 and Proposition 8.1.9 imply that $S_f(t)$ converges to $\frac{1}{2}\big(f(t+)+f(t-)\big)$.

As for the partial sums, using that the Fejér kernel $K_n$ is non-negative and (8.5),

$$|C_{f,n}(t)| \leq \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x-t)|\,K_n(t)\,dt \leq \frac{1}{\pi}F(b)\int_{-\pi}^{\pi} K_n(t)\,dt = F(b).$$

We also have, using again Lemma 8.1.6

$$|S_{f,n}(t) - C_{f,n}(t)| = \frac{1}{n+1}\sum_{k=0}^{n}\big(S_{f,n}(t) - S_{f,k}(t)\big) = \frac{1}{n+1}\sum_{k=0}^{n-1} k|\hat{f}(k)|$$
$$\leq \frac{1}{n+1}\sum_{k=1}^{n-1} 2\pi F(b) \leq 2\pi F(b)$$

Thus

$$|S_{f,n}(t)| \leq 2\pi F(b) + |C_{f,n}(t)| \leq (2\pi+1)F(b). \qquad \square$$

When $f$ is not of bounded variation as in Corollary 8.1.13, the Fourier series is still decent enough, as seen below. We recall Theorem 2.11.9 about Lebesgue points.

***8.1.14. Theorem. (Lebesgue)***

*Let $f \in L^1[-\pi,\pi]$. Then $C_{f,n}(t_0) \to f(t_0)$ whenever $t_0$ is a Lebesgue point for $f$. In particular $C_{f,n} \to f$ a.e.*

*Proof.* By replacing $f$ with $f - f(t_0)$ we assume without loss of generality that $f(t_0) = 0$. That $t_0$ is a Lebesgue point for $f$ then means that

$$\lim_{t\to 0}\frac{1}{2t}\int_{-t}^{t}|f(t_0+t)| = 0.$$

We have, using (8.7) and that $t_0$ is a Lebesgue point for $f$,

$$\int_{|t|<\frac{1}{n}}^{\pi}|f(t_0-t)|\,K_n(t)\,dt \leq \frac{1}{1/n}\int_{|t|\leq\frac{1}{n}}|f(t_0-t)|\,dt \xrightarrow[n\to\infty]{} 0. \tag{8.8}$$

Also, using that $\frac{|t|}{\pi} \leq \sin\frac{|t|}{2}$ on $[-\pi,\pi]$ we can improve (8.6) to

$$K_n(t) \leq \frac{2\pi}{(n+1)t^2}. \tag{8.9}$$

Let $G(t) = \int_0^t \big[|f(t_0+t)| + |f(t_0-s)|\big]\,ds$. The function $G$ is differentiable with derivative $|f(t_0-t|)|$ by the Fundamental Theorem of Calculus (Corollary 2.11.14). This allows us to integrate by parts below. Using that $K_n(-t) = K_n(t)$,

$$\int_{\pi\geq|t|\geq\frac{1}{n}}|f(t_0-t)|\,K_n(t)\,dt = \int_{1/n}^{\pi}\big[|f(t_0+t)| + |f(t_0-t)|\big]\,K_n(t)\,dt$$
$$\leq \frac{2\pi}{(n+1)}\int_{1/n}^{\pi}\frac{|f(t_0+t)| + |f(t_0-t)|}{t^2}\,dt$$

$$= \frac{2\pi}{(n+1)} \frac{G(t)}{t^2}\bigg|_{1/n}^{\pi} + \frac{4\pi}{(n+1)} \int_{1/n}^{\pi} \frac{G(t)}{t^3}\, dt$$

$$= \frac{2\pi G(\pi)}{(n+1)} - \frac{2\pi n^2 G(1/n)}{n+1} + \frac{4\pi}{(n+1)} \int_{1/n}^{\pi} \frac{G(t)}{t^3}\, dt.$$

Of the three terms above, the first one clearly goes to zero with $n$. The second can be written as

$$\frac{2\pi n}{n+1}\, \frac{G(1/n)}{1/n} \xrightarrow[n\to\infty]{} 0$$

since $t_0$ is a Lebesgue point for $f$. As for the third term, fix $\varepsilon > 0$ and choose $\delta > 0$ such that $\frac{G(t)}{t} < \varepsilon$ when $0 < t < \delta$. Then

$$\begin{aligned}
\frac{4\pi}{(n+1)} \int_{1/n}^{\pi} \frac{G(t)}{t^3}\, dt &= \frac{4\pi}{(n+1)} \int_{1/n}^{\delta} \frac{G(t)}{t^3}\, dt + \frac{4\pi}{(n+1)} \int_{\delta}^{\pi} \frac{G(t)}{t^3}\, dt \\
&\le \frac{4\pi}{(n+1)} \int_{1/n}^{\delta} \frac{\varepsilon}{t^2}\, dt + \frac{4\pi}{\delta^3(n+1)} \int_{\delta}^{\pi} G(t)\, dt \\
&\le \frac{4\pi n\varepsilon}{(n+1)} \int_{1/n}^{\delta} \frac{\varepsilon}{t^2}\, dt + \frac{4\pi}{\delta^3(n+1)} \int_{\delta}^{\pi} G(t)\, dt \\
&\le 4\pi\varepsilon + \frac{4\pi}{\delta^3(n+1)} \int_{\delta}^{\pi} G(t)\, dt \xrightarrow[n\to\infty]{} 4\pi\varepsilon.
\end{aligned}$$

This shows that the limsup of the third term is less than $\varepsilon$, and as $\varepsilon$ was arbitrary the limit is zero.

Hence

$$\int_{\pi\ge|t|\ge\frac{1}{n}} |f(t_0 - t)|\, K_n(t)\, dt \xrightarrow[n\to\infty]{} 0. \tag{8.10}$$

Collecting the estimates in (8.8) and (8.10),

$$\begin{aligned}
|C_{f,n}(t_0)| &= \left| \frac{1}{\pi} \int_{-\pi}^{\pi} f(t_0 - t)\, K_n(t)\, dt \right| \le \frac{1}{\pi} \int_{-\pi}^{\pi} |f(t_0 - t)|\, K_n(t)\, dt \\
&= \int_{|t|<\frac{1}{n}}^{\pi} |f(t_0 - t)|\, K_n(t)\, dt + \int_{\pi\ge|t|\ge\frac{1}{n}} |f(t_0 - t)|\, K_n(t)\, dt \\
&\xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

□

*8.1.15. Remark*. There is not much room to relax the hypotheses of Proposition 8.1.4 and still have $S_f(t)$ converging to $t$ everywhere. Concretely, for each $t \in [-\pi, \pi]$ there exists $f \in C[-\pi, \pi]$ such that $S_f(t)$ diverges. To see this, we use the fact that $\|D_n\|_1 \to \infty$ (Exercise 8.1.8). Let $\gamma_n : C[-\pi, \pi] \to \mathbb{C}$ be given by $\gamma_n(f) = S_{f,n}(0)$. We have

$$\gamma_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, D_n(t)\, dt.$$

As $D_n \in L^1[-\pi,\pi]$ (since it is bounded by $n+\frac{1}{2}$ and the interval is finite), Hölder's inequality implies that $\gamma_n$ is a bounded linear functional on $C[-\pi,\pi]$ induced by $D_n$. We know from Riesz–Markov (Theorem 5.6.15) that $\|\gamma_n\|$ is equal the total variation of the measure induced by $D_n$. By Proposition 2.10.12, $\|\gamma_n\| = \|D_n\|_1$. Since the set $\{\|\gamma_n\|\}$ is unbounded, the Uniform Boundedness Principle (Theorem 6.3.16) implies that there exists $f \in C[-\pi,\pi]$ such that $\{\gamma_n(f)\}$ is unbounded. That is, $\{S_{f,n}(0)\}$ is unbounded.

There was nothing particular about $t=0$ in the argument above, as one can define $\gamma_n$ at any other point.

While the existence of $f$ above is very non-explicit (and it shows the power of Functional Analysis!), it is possible to give explicit examples of continuous functions with divergent Fourier series (Exercise 8.1.12).

**8.1.3. Exercises.**

(8.1.1) Let $f(t) = t$, $t \in [-\pi,\pi]$. Show that

$$S_{f,n}(t) = \sum_{k=1}^{n} \frac{(-1)^{k+1}\sin kt}{k}.$$

(8.1.2) Let $f(t) = t^2$, $t \in [-\pi,\pi]$. Show that

$$S_{f,n}(t) = \frac{\pi^2}{3} + \sum_{k=1}^{n} \frac{4(-1)^k \cos kt}{k^2}.$$

(8.1.3) Let $f(t) = e^t$, $t \in [-\pi,\pi]$. Show that

$$S_{f,n}(t) = \frac{\sinh\pi}{2\pi} + \sum_{k=1}^{n} \frac{2(-1)^k \sinh\pi\,\big(\cos kt - k\,\sin kt\big)}{\pi(1+k^2)}.$$

(8.1.4) Let $f : [-\pi,\pi] \to \mathbb{C}$ be given by $f(t) = \operatorname{sgn}(t)$. Show that

$$S_f(t) = \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\sin(2k-1)t}{2k-1}.$$

If available, use graphing software to display $S_{f,n}(t)$ for $n=5$, $n=50$, and $n=500$.

(8.1.5) Mimicking Example 8.1.7, show that

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}.$$

(8.1.6) Let $m, n \in \mathbb{N}$. Show that, as functions,
$$\cos^m t\, \sin^n t \in \operatorname{span}\{a\cos kt + b\sin kt:\ a, b \in \mathbb{R},\ k \in \mathbb{N}\}.$$

(8.1.7) Let
$$f(t) = \begin{cases} \frac{\pi - t}{2}, & 0 \leq t \leq \pi \\ -\frac{\pi+t}{2}, & -\pi \leq t < 0 \end{cases}$$
Use $f$ and Corollary 8.1.13 to show that there exists $c > 0$ such that $\left|\sum_{k=1}^n \frac{\sin kx}{k}\right| \leq c$ for all $n$ and all $x$.

(8.1.8) Show that $\lim_n \|D_n\|_1 = \infty$.

(8.1.9) Let $f \in L^1[-\pi, \pi]$ such that $f$ is even. Show that $S_f(t)$ is of the form
$$S_f(t) = \sum_{k=0}^\infty a_k \cos kt,$$
where
$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt, \qquad a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\cos kt\,dt, \quad k \geq 1.$$

(8.1.10) Let $f \in L^1[-\pi, \pi]$ such that $f$ is odd. Show that $S_f(t)$ is of the form
$$S_f(t) = \sum_{k=1}^\infty b_k \sin kt,$$
where
$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\sin kt\,dt, \qquad\qquad k \geq 1.$$

(8.1.11) Given $f, g \in L^1[-\pi, \pi]$, consider the **convolution**
$$(f * g)(t) = \int_{-\pi}^{\pi} f(x)\,g(t-x)\,dx,$$
where $f, g$ are considered extended periodically to $[-2\pi, 2\pi]$. Show that $\widehat{f * g}(n) = 2\pi\,\hat{f}(n)\,\hat{g}(n)$.

(8.1.12) Use the following steps to show that the function
$$f(t) = \sum_{k=1}^\infty \frac{1}{k^2}\sin\left[\left(2^{k^3}+1\right)\frac{|t|}{2}\right], \qquad\qquad t \in [0, \pi]$$
is in $C[-\pi, \pi]$ and $\{S_{f,n}(0)\}$ diverges.

(i) Show that $f$ is continuous.

(ii) Use Exercise 8.1.9 to write $S_f(t) = \sum_{k=1}^{\infty} a_k \cos kt$, where

$$a_k = \frac{2}{\pi} \sum_{j=1}^{\infty} \frac{1}{j^2} \lambda_{k,2^{j^3-1}},$$

and

$$\lambda_{k,h} = \int_0^{\pi} \sin\left[(2h+1)\frac{t}{2}\right] \cos kt \, dt.$$

(iii) Show that $\sum_{k=0}^{h} \lambda_{k,h} \geq \frac{1}{2} \log h$ and $\sum_{k=0}^{n} \lambda_{k,h} \geq 0$ for all $n, h$.

(iv) Conclude that there exists a sequence $\{n_j\}$ such that

$$S_{f,n_j}(0) \to \infty,$$

by showing that

$$S_{f,2^{j^3}-1}(0) \geq \frac{j^3-1}{2j^2} \log 2.$$

## 8.2. The Fourier Transform

The Fourier transform can be seen as a continuous version of what was done with Fourier series, which was to represent a function in terms of coefficients associated with frequencies. We will consider the several-variable version of the transform, for no technical difficulties arise compared to the one-dimensional case.

**8.2.1. Definition.**

*Let $f \in L^1(\mathbb{R}^n)$. The* ***Fourier Transform*** *of $f$ is the function $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ given by*

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \langle \xi, x\rangle}\, dx, \tag{8.11}$$

*where $dx$ represents Lebesgue measure in $\mathbb{R}^n$. The* ***inverse Fourier transform*** *of a function $f$ is*

$$\check{f}(x) = \int_{\mathbb{R}^n} f(\xi)\, e^{2\pi i \langle x, \xi\rangle}\, d\xi.$$

The integral in the Fourier transform is guaranteed to converge by $f$ being in $L^1$. The inverse Fourier transform is not the inverse of the Fourier transform in general, but it is in a significant number of cases.

*8.2.2. Example*. Here is a classic example in dimension one. Fix $a > 0$ and let $f = 1_{[-a,a]}$. Then

$$\hat{f}(\xi) = \int_{-a}^{a} e^{-2\pi i x\,\xi}\,dx = \frac{\sin a\xi}{\pi\xi}.$$

*8.2.3. Example*. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = e^{-c|x|}$ with $c > 0$. Then

$$\begin{aligned}
\hat{f}(\xi) &= \int_{-\infty}^{\infty} e^{-c|x|}e^{-2\pi i\xi x}\,dx \\
&= \int_{-\infty}^{0} e^{-c|x|}e^{-2\pi i\xi x}\,dx + \int_{0}^{\infty} e^{-c|x|}e^{-2\pi i\xi x}\,dx \\
&= \int_{0}^{\infty} e^{(-c+2\pi i\xi)x}\,dx + \int_{0}^{\infty} e^{-c-2\pi i\xi)x}\,dx \\
&= \frac{1}{c-2\pi i\xi} + \frac{1}{c+2\pi i\xi} = \frac{2c}{c^2+4\pi^2\xi^2}.
\end{aligned}$$

*8.2.4. Example*. Let $\psi(x) = e^{-\pi a\|x\|^2}$ with $a > 0$. Then $\hat{\psi}(x) = \frac{1}{a^{n/2}}e^{-\pi\|x\|^2/a}$. In particular, when $a = 1$ we have $\hat{\psi} = \psi$. Indeed, consider first the one-dimensional case. We have

$$\psi'(x) = -2\pi a x\psi(x).$$

Now take Fourier transform, using Exercises 8.2.2 and 8.2.3, to get

$$2\pi i\xi\,\hat{\psi}(\xi) = -2\pi a\,\frac{(\hat{\psi})'(\xi)}{-2\pi i} = \frac{a}{i}\,(\hat{\psi})'(\xi).$$

Hence

$$\frac{(\hat{\psi})'(\xi)}{\hat{\psi}(\xi)} = -\frac{2\pi\xi}{a}.$$

Thus

$$\hat{\psi}(\xi) = ce^{-\frac{\pi\xi^2}{a}},$$

and $c = \hat\psi(0) = 1/\sqrt{a}$, so $\hat\psi(x) = \frac{1}{\sqrt{a}} e^{-\pi|x|^2/a}$ when $n = 1$. For arbitrary $n$, Fubini as in Theorem 2.7.16 applies, and

$$\begin{aligned}\hat\psi(\xi) &= \int_{\mathbb{R}^n} e^{-\pi a\|x\|^2}\, e^{-2\pi i\langle \xi, x\rangle}\, dx \\ &= \int_{\mathbb{R}^n} e^{-\pi a\sum_j x_j^2}\, e^{-2\pi i\sum_j \xi_j x_j}\, dx_1\cdots dx_n \\ &= \prod_{j=1}^n \int_{\mathbb{R}} e^{-\pi a x_j^2}\, e^{-2\pi i\xi_j x_j}\, dx_j = \prod_{j=1}^n \frac{1}{\sqrt{a}} e^{-\pi a\xi_j^2} \\ &= \frac{1}{a^{n/2}} e^{-\pi\sum_j \xi_j^2/a} = \frac{1}{a^{n/2}}\, e^{-\pi\|\xi\|^2/a}.\end{aligned}$$

**8.2.5. Lemma.**

*Let $f \in L^1(\mathbb{R})$, $y \in \mathbb{R}^n$, and $g \in L^1(\mathbb{R}^n)$ given by $g(x) = f(x-y)$. Then*
$$\hat g(\xi) = e^{-2\pi i\langle \xi, y\rangle}\, \hat f(\xi).$$

*Proof.*
$$\begin{aligned}\hat g(\xi) &= \int_{\mathbb{R}^n} f(x-y)\, e^{-2\pi i\langle \xi, x\rangle}\, dx = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i\langle \xi, x+y\rangle}\, dx \\ &= e^{-2\pi i\langle \xi, y\rangle} \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i\langle \xi, x\rangle}\, dx = e^{-2\pi i\langle \xi, y\rangle}\, \hat f(\xi).\end{aligned}$$
□

**8.2.6. Proposition.**

*Let $f \in L^1(\mathbb{R}^n)$. Then $\hat f$ is uniformly continuous and $\|\hat f\|_\infty \le \|f\|_1$.*

*Proof.* The norm estimate is direct:
$$|\hat f(\xi)| \le \int_{\mathbb{R}^n} |f(x)|\, dx = \|f\|_1,$$
so $\|\hat f\|_\infty \le \|f\|_1$.

Suppose first that $f$ is supported on the ball $B_r(0)$, that is $\|x\| \le r$. The Mean Value Theorem, applied to the function $t \longmapsto e^{-2\pi i t}$, tells us that given $\xi, \eta, x \in \mathbb{R}^n$ there exists $c \in \mathbb{R}$ (depending on all three vectors) such that
$$e^{-2\pi i\langle \xi, x\rangle} - e^{-2\pi i\langle \eta, x\rangle} = -2\pi i\, e^{ic}\, (\langle \xi, x\rangle - \langle \eta, x\rangle).$$

Then

$$\begin{aligned}
|\hat{f}(\xi) - \hat{f}(\eta)| &= \left| \int_{\mathbb{R}^n} f(x) \left( e^{-2\pi i \langle \xi, x \rangle} - e^{-2\pi i \langle \eta, x \rangle} \right) dx \right| \\
&= 2\pi \left| \int_{\mathbb{R}^n} f(x)\, e^{ic} \, \langle \xi - \eta, x \rangle \, dx \right| \\
&\leq 2\pi r \, \|\xi - \eta\| \, \|f\|_1 .
\end{aligned}$$

For arbitrary $f \in L^1(\mathbb{R}^n)$, since $f = \lim_r f\, 1_{B_r(0)}$ by Dominated Convergence we have

$$\|f - f\, 1_{B_r(0)}\|_1 = \int_{\mathbb{R}^n} |f - f\, 1_{B_r(0)}| \, dm \xrightarrow[r \to \infty]{} 0.$$

Fix $\varepsilon > 0$. Let $g = f\, 1_{B_r(0)}$ with $r$ big enough such that $\|f - g\|_1 < \varepsilon$. By the above there exists $\delta > 0$ such that $|\hat{g}(\xi) - \hat{g}(\eta)| < \varepsilon$ if $\|\xi - \eta\| < \delta$. Then

$$|\hat{f}(\xi) - \hat{g}(\xi)| \leq \|f - g\|_1,$$

and

$$\begin{aligned}
|\hat{f}(\xi) - \hat{f}(\eta)| &\leq |\hat{f}(\xi) - \hat{g}(\xi)| + |\hat{g}(\xi) - \hat{g}(\eta)| + |\hat{g}(\eta) - \hat{f}(\eta)| \\
&\leq 2\|f - g\|_1 + \varepsilon < 3\varepsilon.
\end{aligned}$$

Thus $\hat{f}$ is uniformly continuous. □

Though not visible at first sight, the Fourier transform has a nice relation with the 2-norm. Here is a first bit of evidence for it.

***8.2.7. Lemma.***

*Let $f, g \in L^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} g(\xi)\, \hat{f}(\xi)\, d\xi = \int_{\mathbb{R}^n} \hat{g}(x)\, f(x)\, dx.$$

*Proof.* Since $\|\hat{f}\|_1 \leq \|f\|_1$, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(\xi)|\, |f(x)\, e^{-2\pi i \langle x, \xi \rangle}| \, d\xi\, dx &\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(\xi)|\, |f(x)| \, d\xi\, dx \\
&\leq \|g\|_1 \|f\|_1 < \infty.
\end{aligned}$$

Then Fubini's Theorem (as in Theorem 2.7.16) applies and we can exchange integrals. Hence

$$\int_{\mathbb{R}^n} g(\xi)\, \hat{f}(\xi)\, d\xi = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(\xi)\, f(x)\, e^{-2\pi i \langle \xi, x \rangle} \, d\xi\, dx$$

$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(\xi)\, f(x)\, e^{-2\pi i \langle \xi, x\rangle}\, dx\, d\xi$$
$$= \int_{\mathbb{R}^n} \hat{g}(x)\, f(x)\, dx. \qquad \square$$

**8.2.8. Proposition.** *(Inverse Fourier Transform)*

*Let $f \in L^1(\mathbb{R}^n)$ with $\hat{f} \in L^1(\mathbb{R}^n)$. Then $\check{\hat{f}}(x) = f(x) = \hat{\check{f}}(x)$ almost everywhere.*

*Proof.* We will use mollifiers as in Proposition 2.8.22. Let $\psi(x) = e^{-\pi\|x\|^2}$; we know from Example 8.2.4 that $\hat{\psi} = \psi$. Let $\psi_m(x) = m^n\psi(mx)$; a straightforward computation shows that $\psi_m = \widehat{\varphi_m}$, where $\varphi_m(x) = e^{-\pi\|x\|^2/m^2}$. By Corollary 2.11.10, $\lim_m \psi_m * f(x) = f(x)$ almost everywhere. For those $x$ where the equality holds, using Lemma 8.2.5 and Lemma 8.2.7,

$$\begin{aligned}
f(x) = \lim_m \psi_m * f(x) &= \lim_m \int_{\mathbb{R}^n} m^n e^{-\pi m^2\|x-y\|^2}\, f(y)\, dy \\
&= \lim_m \int_{\mathbb{R}^n} \hat{\varphi}(y-x)\, f(y)\, dy = \lim_m \int_{\mathbb{R}^n} \hat{\varphi}(y)\, f(y+x)\, dy \\
&= \lim_m \int_{\mathbb{R}^n} \varphi(y)\, e^{2\pi i\langle y,x\rangle}\, \hat{f}(y)\, dy \\
&= \lim_m m^n \int_{\mathbb{R}^n} \frac{1}{(m^2)^{n/2}} e^{-\pi\|y\|^2/m^2 + 2\pi i\langle y,x\rangle}\, \hat{f}(y)\, dy \\
&= \lim_m \int_{\mathbb{R}^n} e^{-\pi\|y\|^2/m^2 + 2\pi i\langle y,x\rangle}\, \hat{f}(y)\, dy \\
&= \int_{\mathbb{R}^n} e^{2\pi i\langle y,x\rangle}\, \hat{f}(y)\, dy = \check{\hat{f}}(x),
\end{aligned}$$

where the limit in the last integral uses Dominated Convergence. The equality $\hat{\check{f}}(x) = f(x)$ follows by repeating the proof with a version of Lemma 8.2.7 for the inverse Fourier transform. $\square$

Here we can talk briefly about what the Fourier Transform means. Using Proposition 8.2.8 we can write, for $f \in L^1(\mathbb{R})$ and such that $\hat{f} \in L^1(\mathbb{R}^n)$,

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

Compare this with a Fourier series, (8.2). There is a clear formal analogy here. For a Fourier series, $\hat{f}(n)$ is the coefficient that regulates the contribution of the

$n^{\text{th}}$ harmonic to the series for $f$. So it makes sense to interpret the number $\hat{f}(\xi)$ as somehow measuring the contribution of the circular function of frequency $\xi$ to $f$.

In general the Fourier transform of a function in $L^1(\mathbb{R}^n)$ will not be in $L^1(\mathbb{R}^n)$. We will soon prove that functions in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ are mapped into $L^2(\mathbb{R}^n)$; for that we need a bit of preparation.

**8.2.9. Lemma.**

*Let $f \in C^\infty(\mathbb{R}^n) \cap C_c(\mathbb{R}^n)$. Then $\hat{f} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.*

*Proof.* Since $f$ has compact support, for $g \in C^1(\mathbb{R}^n)$ we have the integration by parts formula

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j}\, g = -\int_{\mathbb{R}^n} f\, \frac{\partial g}{\partial x_j}.$$

Then

$$\begin{aligned}\widehat{\frac{\partial f}{\partial x_j}}(\xi) &= \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j}(y)\, e^{-2\pi i\langle \xi, y\rangle}\, dy \\ &= -\int_{\mathbb{R}^n} f(y)\,(-2\pi i\xi_j)\, e^{-2\pi i\langle \xi, y\rangle}\, dy = 2\pi i\xi_j\, \hat{f}(\xi).\end{aligned}$$

Thus

$$\widehat{\frac{\partial^n f}{\partial x_1\cdots\partial x_n}}(\xi) = (-1)^n(2\pi)^{2n}\xi_1\cdots\xi_n\, \hat{f}(\xi).$$

The left-hand-side, being the Fourier transform of a function in $L^1(\mathbb{R}^n)$ (since it is continuous with compact support) is bounded. So there exists $c > 0$ such that

$$|\hat{f}(\xi)| \leq \frac{c}{|\xi_1\cdots\xi_n|}.$$

Then

$$\int_{\|\xi\|_\infty\geq 1} |\hat{f}(\xi)|^2\, d\xi \leq c\int_{\|\xi\|_\infty\geq 1} \frac{1}{\xi_1^2\cdots\xi_n^2}\, d\xi = c\prod_{j=1}^n \int_1^\infty \frac{1}{t^2}\, dt < \infty.$$

Hence $\hat{f} \in L^2(\mathbb{R}^n)$, as $\hat{f}$ is bounded and hence integral on $\|\xi\|_\infty \leq 1$. Repeating the argument while differentiating twice in each variable we can get the estimate $|\hat{f}(\xi)| \leq \frac{c}{\xi_1^2\cdots\xi_n^2}$, and it follows that $\hat{f} \in L^1(\mathbb{R}^n)$. □

**8.2.10. Proposition. *(Plancherel)***

*Let $f, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$. In particular, $\|\hat{f}\|_2 = \|f\|_2$.*

*Proof.* For a function $h \in L^1(\mathbb{R}^n)$, if $\bar{h}$ denotes the function $x \longmapsto \overline{h(x)}$, we have $\check{\bar{h}}(x) = \bar{\hat{h}}(x)$. Indeed, since taking conjugate is continuous,

$$\check{\bar{h}}(x) = \int_{\mathbb{R}^n} \overline{h(\xi)}\, e^{2\pi i \langle \xi, x\rangle}\, d\xi = \overline{\int_{\mathbb{R}^n} h(\xi)\, e^{-2\pi i \langle \xi, x\rangle}\, d\xi} = \overline{\hat{h}(x)} = \bar{\hat{h}}(x).$$

Suppose first that $f, g \in C^\infty(\mathbb{R}^n) \cap C_c(\mathbb{R}^n)$. Then $\hat{f}, \hat{g} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ by Lemma 8.2.9. Using this and Lemma 8.2.7 and Proposition 8.2.8,

$$\begin{aligned}\langle \hat{f}, \hat{g}\rangle &= \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{g}(\xi)}\, d\xi = \int_{\mathbb{R}^n} f(x)\, \hat{\bar{\hat{g}}}(x)\, dx = \int_{\mathbb{R}^n} f(x)\, \bar{\check{\hat{g}}}(x)\, dx \\ &= \int_{\mathbb{R}^n} f(x)\, \overline{g(x)}\, dx = \langle f, g\rangle.\end{aligned}$$

If we denote by $\mathcal{F} : L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ the Fourier transform, we have shown that it is a linear isometric operator from $C^\infty(\mathbb{R}^n) \cap C_c(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$. By Proposition 2.8.22, the space $C^\infty(\mathbb{R}^n) \cap C_c(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, and $\mathcal{F}$ admits a unique extension by Proposition 6.1.9. The continuity of the norm guarantees that the extension—that we will denote $\mathcal{F}$—is still isometric. Thus $\|\hat{f}\|_2 = \|f\|_2$ for all $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. The Polarization Identity (4.6) then gives $\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$ for all $f, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. $\square$

Proposition 8.2.10 shows that the Fourier transform is a linear isometric operator $\mathcal{F} : L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ . Its image consists of continuous functions by Proposition 8.2.6. Given $g \in C^\infty(\mathbb{R}^n) \cap C_c(\mathbb{R}^n)$, by Lemma 8.2.9, applied to the inverse Fourier transform, $\hat{g} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then Proposition 8.2.8 shows that $\mathcal{F}(\hat{g}) = g$ as elements of $L^2(\mathbb{R}^n)$. Thus $\mathcal{F}$ maps $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ into a dense subspace of $L^2(\mathbb{R}^n)$. By Proposition 6.1.9 we get a bounded extension to $L^2(\mathbb{R}^n)$, which will still be isometric by the continuity of the norm. As $\mathcal{F}$ is isometric with dense range, its image is closed (Exercise 6.1.9) and hence all of $L^2(\mathbb{R}^n)$; so $\mathcal{F}$ is an isometric linear isomorphism of $L^2(\mathbb{R}^n)$ onto itself. This way the Fourier transform is defined for all $f \in L^2(\mathbb{R}^n)$, though $\hat{f}$ will not be given by (8.11) when $f$ is not in $L^1(\mathbb{R}^n)$.

*8.2.11. Example.* Suppose we want to find the Fourier transform of $f(x) = \frac{1}{1+x^2}$. The integral that defines $\hat{f}$ does not look very friendly. On the other hand, we know from Example 8.2.3 that $f = \hat{g}$, where $g(x) = \pi e^{-2\pi|x|}$. Noting that $\hat{f}(-\xi) = \check{f}(\xi)$, we have

$$\hat{f}(\xi) = \check{f}(-\xi) = \check{\hat{g}}(-\xi) = g(-\xi) = \pi e^{-2\pi|\xi|}.$$

**8.2.1. Exercises.**

(8.2.1) Let $f \in L^1(\mathbb{R}^n)$ and $c \in \mathbb{R}$. Show that the Fourier transform is linear, and that if $g(x) = f(cx)$, then $\hat{g}(\xi) = \frac{1}{|c|^n}\,\hat{f}\left(\frac{\xi}{c}\right)$.

(8.2.2) Let $f \in L^1(\mathbb{R})$ and such that $g : x \longmapsto xf(x)$ is also in $L^1$. Show that $\hat{f}$ is differentiable and $(\hat{f})'(\xi) = -2\pi i \hat{g}(\xi)$.

(8.2.3) Let $f \in L^1(\mathbb{R})$ differentiable with $f' \in L^1(\mathbb{R})$. Show that $\widehat{f'}(\xi) = 2\pi i\xi\,\hat{f}(\xi)$.

(8.2.4) Let $f, g \in L^1(\mathbb{R}^n)$. Show that

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\,\hat{g}(\xi).$$

## 8.3. Hilbert Function Spaces

**8.3.1. Definition.**

*A* ***Hilbert Function Space*** *is a Hilbert space $\mathcal{H}$ that consists of functions over a set $X$, and such that all point evaluations are bounded.*

*A* ***kernel*** *is a function $k : X \times X \to \mathbb{C}$ such that there is a map $\varphi_k : X \to \mathcal{H}$ such that*

$$k(x, y) = \langle \varphi_k(x), \varphi_k(y) \rangle.$$

*The map $\varphi_k$ is known as a* ***feature map*** *for $k$.*

This definition is a bit at odds with previous claims that Hilbert spaces do not really depend on their concrete representation. And that is true, in everything we have done with Hilbert spaces so far. But we will see here cases where one can gain an advantage of using different presentations of the same Hilbert space.

Hilbert function spaces are sometimes called **Reproducing Kernel Hilbert Spaces**.

Consider a Hilbert Space Function $\mathcal{H}$ over $X$, and let $x \in X$. By definition, the linear map $f \longmapsto f(x)$ is bounded. Then the Riesz Representation Theorem (4.5.4) gives us $k_x \in \mathcal{H}$ such that $f(x) = \langle f, k_x \rangle$ for all $f \in \mathcal{H}$. Since $k_x \in \mathcal{H}$, this gives us

$$k_x(y) = \langle k_x, k_y \rangle. \tag{8.12}$$

Thus $k(x, y) = k_x(y)$ is a kernel, with feature map $\varphi_k(x) = k_x$.

*8.3.2. Examples.*

(a) Let $\mathcal{H} = \mathbb{C}^n$ with the usual inner product. We can see $\mathcal{H}$ as $\ell^2(X)$, where $X = \{1, \dots, n\}$. As $X$ is finite, this is simply the set of all functions $X \to \mathbb{C}$. Given any $f \in \mathcal{H}$ and $x \in X$, we have $f(x) = \langle f, e_x \rangle$, where $e_1, \dots, e_n$ is the canonical basis. Thus we have the kernel

$$k(x, y) = \langle e_x, e_y \rangle = \delta_{x,y}$$

and the feature map $x \longmapsto e_x$.

(b) Let $\mathcal{H} = \ell^2(\mathbb{N})$. The exact same idea applies, now with $X = \mathbb{N}$, and $k(x, y) = \delta_{x,y}$ for $x, y \in \mathbb{N}$ and feature map $x \longmapsto e_x$.

(c) Let $\mathcal{H} = L^2[0, 1]$. This is *not* a Hilbert Function Space, as its elements are classes of functions, and point evaluations make no sense.

(d) The **Hardy space** $\mathbb{H}^2(\mathbb{D})$ is the space of analytic functions $\sum_n a_n z^n$ such that $a \in \ell^2(\mathbb{N})$, with the inner product $\langle f, g \rangle = \sum_n a_n \overline{b_n}$ if $f = \sum_n a_n z^n$ and $g = \sum_n b_n z^n$. It is by construction canonically isomorphic to $\ell^2(\mathbb{N})$, so it is a Hilbert space. Formally, one needs to check that if $a \in \ell^2(\mathbb{N})$ then $f = \sum_n a_n z_n$ is analytic on $\mathbb{D}$ (Exercise 8.3.1). The evaluation map $f \longmapsto f(z)$ is bounded, for

$$|f(z)|^2 = \Big| \sum_n a_n z^n \Big|^2 \le \sum_n |a_n|^2 \sum_n |z|^{2n} = \frac{\|a\|^2}{1 - |z|^2}.$$

For any $f \in H^2(\mathbb{D})$ we have $f(z) = \sum_n a_n z^n = \langle f, k_z \rangle$, where

$$k_z(w) = \sum_{n=0}^{\infty} \bar{z}^n \, w^n = \frac{1}{1 - \bar{z}w} = k(z, w).$$

Thus $k_z$ is our feature map, and $k(z, w) = \sum_n z^n \overline{w^n} = \frac{1}{1 - z\overline{w}}$. This is known as the **Szego kernel**.

(e) The **Bergman space** is

$$L^2_a(\mathbb{D}) = \{f : \mathbb{D} \to \mathbb{C}, \text{ holomorphic}, \int_{\mathbb{D}} |f|^2 \, dm < \infty\},$$

with the inner product

$$\langle f, g \rangle_B = \frac{1}{\pi} \int_{\mathbb{D}} f(z) \, \overline{g(z)} \, dm(z).$$

A direct computation (Exercise 8.3.2) shows that if $f(z) = \sum_{n=0}^{\infty} a_n z^n$, $\sum_{n=0}^{\infty} b_n z^n$, then

$$\langle f, g \rangle_B = \sum_{n=0}^{\infty} \frac{a_n \, \overline{b_n}}{n + 1}.$$

We leave as an exercise to show that the Bergman space is a function space (that is, it is complete and point evaluations are bounded), and that the set $\{z^n\}_{n=0}^{\infty}$ is orthogonal and complete (Exercise 8.3.3). From the formula for the inner product above we see that

$$k_z(w) = \sum_{n=0}^{\infty}(n+1)\bar{z}^n\, w^n = \frac{1}{(1-\bar{z}\,w)^2} = k(z,w).$$

(f) The **Dirichlet space** $\mathcal{D}(\mathbb{D})$ is the space of analytic functions $f$ on $\mathbb{D}$ such that the integral $\int_{\mathbb{D}} |f'|\, dm$ is finite. One considers the inner product

$$\langle f, g\rangle = f(0)\overline{g(0)} + \frac{1}{\pi}\int_{\mathbb{D}} f'(z)\,\overline{g'(z)}\, dz.$$

We leave it as an exercise to show that $\mathcal{D}(\mathbb{D})$ is a Hilbert function space, and to find its reproducing kernel and its feature map (Exercise 8.3.5).

(g) Fix $a > 0$. Let $\mathcal{H} = \{f \in C(\mathbb{R}) \cap L^2(\mathbb{R}) : \ \operatorname{supp}\hat{f} \subset [-a,a]\}$, as a subspace of $L^2(\mathbb{R})$. It is complete, for if $\{f_n\} \subset \mathcal{H}$ and $\|f_n - g\|_2 \to 0$, then $\|\hat{f}_n - \hat{g}\|_2 \to 0$ (Proposition 8.2.10). The support cannot be enlarged by taking limit, so $\operatorname{supp}\hat{g} \subset [-a,a]$. And, by Proposition 8.2.8 (which applies because $\hat{g}$ is in $L^{\infty}(\mathbb{R})$ and has finite support, so it is in $L^1(\mathbb{R})$) we have that $g = \check{\hat{g}}$ almost everywhere, and an inverse Fourier transform of $\hat{g} \in L^2[-a,a] \subset L^1[-a,a]$ is continuous almost everywhere. Hence $\mathcal{H}$ is complete. For almost all $x$ we have, using Cauchy–Schwarz and Plancherel,

$$|f(x)| = \left|\int_{\mathbb{R}} \hat{f}(\xi)\, e^{2\pi i\xi x}\, d\xi\right| \le \|\hat{f}\|_2 \left(\int_{-a}^{a} |e^{2\pi i\xi x}|^2\, d\xi\right)^{1/2} = \sqrt{2a}\,\|f\|_2.$$

As $f$ is continuous, the inequality holds for all $x$, and so all point evaluations are continuous and $\mathcal{H}$ is a Hilbert Function space. Using Fubini as in Theorem 2.7.16,

$$\begin{aligned} f(x) &= \int_{-a}^{a} \hat{f}(\xi)\, e^{2\pi i\xi x}\, d\xi = \int_{-a}^{a}\int_{\mathbb{R}} f(y)\, e^{-2\pi i\xi(y-x)}\, dy\, d\xi \\ &= \int_{\mathbb{R}}\int_{-a}^{a} f(y)\, e^{-2\pi i\xi(y-x)}\, d\xi\, dy = \int_{\mathbb{R}} f(y)\,\frac{\sin 2\pi a(y-x)}{2\pi(y-x)}\, dy. \end{aligned}$$

Hence the reproducing kernel is

$$k(x,y) = \frac{\sin 2\pi a(y-x)}{2\pi(y-x)}.$$

A natural question is which functions $k : X \times X \to \mathbb{C}$ can be reproducing kernels.

**8.3.3. Definition.**

*Let $X$ be a set. A function $k : X \times X \to \mathbb{C}$ is* ***positive definite*** *if for any $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{C}$, $x_1, \ldots, x_n \in X$,*

$$\sum_{r,s=1}^{n} a_r \overline{a_s}\, k(x_r, x_s) \geq 0, .$$

*and the expression can only be zero if $a_1 = \cdots = a_n = 0$.*

***8.3.4. Theorem. (Moore–Aronszajn)***

*Let $X$ be a set, and $k : X \times X \to \mathbb{C}$ a function. The following statements are equivalent:*

*(i) $k$ is positive definite;*

*(ii) there exists a unique Hilbert function space $\mathcal{H}$ such that $k$ is a kernel for $\mathcal{H}$.*

*Proof.* (i) $\implies$ (ii) For each $x \in X$ we define $k_x(y) = k(x, y)$. Let

$$\mathcal{H}_0 = \operatorname{span}\{k_x : \ x \in X\}.$$

A consequence of the hypothesis is that the set $\{k_x : \ x \in X\}$ is linearly independent, so each element in $\mathcal{H}_0$ can be written as a linear combination in a unique way. This allows us to define on $\mathcal{H}_0$ the inner product

$$\Big\langle \sum_{r=1}^{n} a_r k_{x_r}, \sum_{s=1}^{m} b_s k_{x_s} \Big\rangle = \sum_{r=1}^{n} \sum_{s=1}^{m} a_r \overline{b_s}\, k(x_r, x_s).$$

This is positive definite by hypothesis, and sesquilinear by construction. We note that if $f = \sum_{r=1}^{n} a_r k_{x_r}$, then

$$f(x) = \sum_{r=1}^{n} a_r\, k_{x_r}(x) = \sum_{r=1}^{n} a_r\, k(x_r, x) = \langle f, k_x \rangle.$$

Let $\overline{\mathcal{H}_0}$ be the completion of $\mathcal{H}_0$. For $g \in \overline{\mathcal{H}_0}$, we define $\hat{g} : X \to \mathbb{C}$ by $\hat{g}(x) = \langle g, k_x \rangle$. Let $\mathcal{H} = \{\hat{g} : \ g \in \overline{\mathcal{H}_0}\}$ and $\psi : \overline{\mathcal{H}_0} \to \mathcal{H}$ be the map $\psi(g) = \hat{g}$. It is automatic that $\psi$ is linear. Also, if $g \in \mathcal{H}_0$ then $\psi(g)(x) = \hat{g}(x) = g(x)$. If $\psi(g) = 0$, then $\langle g, k_x \rangle = 0$ for all $x \in X$, so $g \in \mathcal{H}_0^{\perp} = \{0\}$; thus $\psi$ is injective. Using that $\psi$ is a linear bijection and defining $\langle \hat{g}, \hat{h} \rangle = \langle g, h \rangle$, we get that $\mathcal{H}$ is a Hilbert space of functions; the evaluation maps are bounded by construction, for $|\hat{g}(x)| = |\langle g, k_x \rangle| \leq \|g\|\, \|k_x\|$. Thus $\mathcal{H}$ is a Hilbert function space, and $k$ is its reproducing kernel since $\hat{g}(x) = \langle g, k_x \rangle$ for all $\hat{g} \in \mathcal{H}$.

(ii) $\implies$ (i) If $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{C}$, $x_1, \ldots, x_n \in X$,

$$\sum_{r,s=1}^{n} a_r \overline{a_s}\, k(x_r, x_s) = \sum_{r,s=1}^{n} a_r \overline{a_s}\, \langle k_r, k_s \rangle = \Big\langle \sum_{r=1}^{n} a_r k_r, \sum_{r=1}^{n} a_r k_r \Big\rangle \geq 0. \qquad \square$$

**8.3.1. Exercises.**

(8.3.1) Show that if $a \in \ell^2(\mathbb{N})$ then $f = \sum_n a_n z^n$ is analytic on $\mathbb{D}$.

(8.3.2) Show that if $f, g \in B_a(\mathbb{D})$ and $f = \sum_n a_n z^n$, $g = \sum_n b_n z^n$, then

$$\langle f, g \rangle_B = \sum_{n=0}^{\infty} \frac{a_n \overline{b_n}}{n+1}.$$

(8.3.3) Show that the Bergman space is a Hilbert function space (that is, complete and point evaluations are bounded), and that the set $\{z^n\}_{n=0}^{\infty}$ is orthogonal and total.

(8.3.4) Let $p \in [1, \infty]$, and consider the corresponding Hardy space $\mathbb{H}^p(\mathbb{D})$, that is the space of analytic functions $f(z) = \sum_n a_n z^n$ on the disk, such that $a \in \ell^p(\mathbb{N})$ and $\sup_{0<r<1} \|f_r\|_p < \infty$, where $f_r(t) = f(re^{it})$, $t \in [-\pi, \pi]$. Show that $\|f\| = \lim_{r \to 1} \|f_r\|_p$ is a norm, and that $\mathbb{H}^p(\mathbb{D})$ is complete with that norm.

(8.3.5) Show that $\mathcal{D}(\mathbb{D})$ is a Hilbert function space, and find its reproducing kernel and its feature map.

(8.3.6) Let $\mathcal{H}$ be a Hilbert Function Space over $X$ with reproducing kernel $k$, and $\mathcal{H}_0 \subset \mathcal{H}$ a closed subspace. Show that $\mathcal{H}_0$ is a Hilbert Function Space over $X$ and that its reproducing kernel is given by $k_0(x, y) = \langle P(k_x), P(k_y) \rangle$, where $P$ is the orthogonal projection onto $\mathcal{H}_0$.

CHAPTER

# 9

# Bounded operators on a Banach space

Throughout this chapter, $\mathcal{X}$ and $\mathcal{Y}$ will be Banach spaces. The main object in this chapter is the set $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ of all bounded linear operators $\mathcal{X} \to \mathcal{Y}$; often times, with $\mathcal{Y} = \mathcal{X}$, in which case we write $\mathcal{B}(\mathcal{X})$.

Linear operators are often studied on their own merits, but sometimes it pays off to study $\mathcal{B}(\mathcal{X}, \mathcal{Y})$, as we have seen in previous chapters from the study of $\mathcal{X}^* = \mathcal{B}(\mathcal{X}, \mathbb{C})$.

## 9.1. Linear operators and continuity

**9.1.1. Bounded operators.** We discussed bounded operators briefly in Section 6.1. In particular we proved, in Proposition 6.1.2, that if $T : \mathcal{X} \to \mathcal{Y}$ is linear, then the following statements are equivalent:

(i) $T$ is continuous;

(ii) $T$ is continuous at some point;

(iii) $T$ is continuous at 0;

(iv) there exists $k > 0$ such that $\|Tx\| \le k\|x\|$ for all $x \in \mathcal{X}$.

This allowed us to define a norm on $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ by

$$\|T\| = \sup\{\|Tx\| : \ x \in \mathcal{X}, \ \|x\| = 1\}.$$

In our initial discussion about bounded operators, we also proved that the space $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is always complete and non-empty; it contains at the very least lots of compact operators. We also showed in Proposition 6.1.9 that a bounded linear operator on a dense subset extends to the closure with the same norm, as long as the codomain is complete.

The set $\mathcal{B}(\mathcal{X}) = \mathcal{B}(\mathcal{X}, \mathcal{X})$ carries a natural structure of an algebra, where the sum is defined pointwise, and the product is given by composition. The operator norm gives $\mathcal{B}(\mathcal{X})$ a canonical structure as a metric space, with the distance $\operatorname{dist}(T, S) = \|S - T\|$.

***9.1.1. Proposition.***

*The assignment $T \longmapsto \|T\|$, from $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ into $\mathbb{R}^+$, is a norm. Moreover, when $\mathcal{Y} = \mathcal{X}$ and $T, S \in \mathcal{B}(\mathcal{X})$, $\|TS\| \leq \|T\| \, \|S\|$.*

*Proof.* This is in Proposition 6.1.4. □

By Proposition 6.1.5, $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is Banach when $\mathcal{Y}$ is Banach. As explained in Section 6.1, the set $\mathcal{B}(\mathcal{X})$ carries at least many finite-rank operators and their limits , and also the identity operator $I_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$, where $I_{\mathcal{X}} x = x$ for all $x \in \mathcal{X}$.

***9.1.2. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{X} \to \mathcal{Y}$ linear. If $T$ is weak-weak continuous, then $T$ is bounded.*

*Proof.* Suppose that $T$ is not bounded. Then there exists a sequence $\{x_n\} \subset \mathcal{X}$ with $x_n \to 0$ and $\|Tx_n\| \to \infty$. Since it is norm convergent, the sequence $\{x_n\}$ converges weakly to 0. Then $Tx_n \xrightarrow{\text{weak}} 0$. By Proposition 7.1.11, the sequence $\{Tx_n\}$ is bounded, a contradiction. □

Note that Proposition 9.1.2 proves a little more than it says, since it does not use that $T$ is weakly continuous, but only that it is sequentially weakly continuous.

**9.1.2. Operator-Block Form.** It is a fairly common situation to have a Banach space $\mathcal{X}$ with a complemented subspace $\mathcal{X}_1$. If $P_1$ is a bounded projection onto $\mathcal{X}_1$ and $P_2 = I_{\mathcal{X}} - P_1$, $\mathcal{X}_2 = P_2\mathcal{X}$, we decompose as $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. In such a case, one can think of $T \in \mathcal{B}(\mathcal{X})$ as a $2 \times 2$ "operator-block matrix"

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}, \tag{9.1}$$

where $T_{11} \in \mathcal{B}(\mathcal{X})$, $T_{12} \in \mathcal{B}(\mathcal{X}_2, \mathcal{X}_1)$, $T_{21} \in \mathcal{B}(\mathcal{X}_1, \mathcal{X}_2)$, $T_{22} \in \mathcal{B}(\mathcal{X}_2)$. Formally, $T_{kj} = P_k T P_j$, interpreted as a linear operator from $\mathcal{X}_j \to \mathcal{X}_k$.; or equivalently, $T_{kj} = P_k T|_{\mathcal{X}_j}$. We will only consider here the $2 \times 2$ case, but everything we say generalizes straightforwardly to the situation where $\mathcal{X}$ is a direct sum of $n$ subspaces and we see $T$ as an $n \times n$ operator-block matrix. Properly, if $P_1, P_2$ are the corresponding projections onto $\mathcal{X}_1, \mathcal{X}_2$, we have

$$Tx = T_{11}P_1x + T_{12}P_2x + T_{21}P_1x + T_{22}P_2x.$$

But this is perfectly coherent with being interpreted as

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} P_1x \\ P_2x \end{bmatrix} = \begin{bmatrix} T_{11}P_1x + T_{12}P_2x \\ T_{21}P_1x + T_{22}P_2x \end{bmatrix},$$

which behaves exactly as the usual "row times column" of scalar matrices. And the same happens with composition of operators:

$$ST = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} S_{11}T_{11} + S_{12}T_{21} & S_{11}T_{12} + S_{12}T_{22} \\ S_{21}T_{11} + S_{22}T_{21} & S_{21}T_{12} + S_{22}T_{22} \end{bmatrix}. \tag{9.2}$$

All the compositions make sense (as confirmed by the inner indices agreeing, we always have compositions of the form $S_{kj}T_{jh}$), and in fact this is precisely the operator-block form of $ST$ (Exercise 9.1.7). In this context, the projections $P_1$ and $P_2$ become

$$P_1 = \begin{bmatrix} I_{\mathcal{X}_1} & 0 \\ 0 & 0 \end{bmatrix}, \qquad P_2 = \begin{bmatrix} 0 & 0 \\ 0 & I_{\mathcal{X}_2} \end{bmatrix}.$$

Let us emphasize that the operator-block matrix representation depends on the choice of the projection $P_1$ and not just on the subspace $\mathcal{X}_1$. Here is an example. Let $\mathcal{X} = \mathbb{C}^2$, $\mathcal{X}_1 = \mathbb{C} \oplus 0$, $\mathcal{X}_2 = 0 \oplus \mathcal{X}_2$. Let $T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. From the above point of view, this presentation is the one that comes from taking $P_1(x, y) = x$. Consider instead the projection $Q_1 : \mathcal{X} \to \mathcal{X}_1$, given by $Q_1(x, y) = (x + 2y, 0)$. Now $Q_2(x, y) = (-2y, y)$. In this presentation the block form of the operator $T$ is

$$T = \begin{bmatrix} 4 & T_{12} \\ T_{21} & -2 \end{bmatrix}.$$

Indeed,

$$\begin{aligned} T_{11}(x, y) = Q_1TQ_1(x, y) &= Q_1T(x + 2y, 0) = Q_1(x + 2y, 3x + 6y) \\ &= (4x + 8y, 0) = 4Q_1(x, y). \end{aligned}$$

Similarly we can check that $T_{22}(-2y, y) = (4y, -2y) = -2(-2y, y)$. The operators $T_{12}$ and $T_{21}$ cannot be associated with numbers unless we make an identification between the ranges of $Q_1$ and $Q_2$ (which is what we usually do when working with the canonical basis). Here

$$T_{12}(x, y) = Q_1TQ_2(x, y) = Q_1T(-2y, y) = Q_1(0, -2y) = (-2y, 0),$$

and $T_{21}(x, y) = 3(-2x - 4y, x + 2y)$.

The operator-block matrix form has many uses, and it will appear from time to time. As an example,

**9.1.3. Proposition.**

*Let $\mathcal{X}$ with complemented subspace $\mathcal{X}_1$, and $T \in \mathcal{B}(\mathcal{X})$. Then $\mathcal{X}_1$ is invariant for $T$ (that is, $T\mathcal{X}_1 \subset \mathcal{X}_1$) if and only if $T_{21} = 0$. That is,*

$$T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$$

*Proof.* Exercise 9.1.8. □

We include here a result that will be needed in Chapter 13.

**9.1.4. Proposition.**

*Let $\mathcal{X}$ be a Banach space and $\mathcal{C} \subset \mathcal{B}(\mathcal{X})$ a convex set. Then the pointwise-weak and pointwise-norm closures of $\mathcal{C}$ agree. That is, if $T \in \mathcal{B}(\mathcal{X})$ and there exists a net $\{T_j\} \subset \mathcal{C}$ with $T_j(x) \xrightarrow{\text{weak}} T(x)$ for all $x \in \mathcal{X}$, then there exists a net $\{T'_F\} \subset \mathcal{C}$ with $T'_F(x) \to T(x)$ for all $x \in \mathcal{X}$.*

*Proof.* As norm-convergence implies weak-convergence, one of the inclusions is trivial. Fix $T \in \mathcal{B}(\mathcal{X})$ such that there exists a net $\{T_j\} \subset \mathcal{C}$ and $T_j(x) \xrightarrow{\text{weak}} T(x)$ for all $x \in \mathcal{X}$. That is, $\varphi(T_j(x)) \to \varphi(T(x))$ for all $\varphi \in \mathcal{X}^*$ and all $x \in \mathcal{X}$. Fix $F = \{x_1, \ldots, x_n\} \subset \mathcal{X}$. We consider the maps

$$T_j^{(n)} : \left(\bigoplus_{r=1}^{n} \mathcal{X}\right)_{\ell^1} \to \left(\bigoplus_{r=1}^{n} \mathcal{X}\right)_{\ell^1}$$

given by $T_j^{(n)}(z_1, \ldots, z_n) = (T_j z_1, \ldots, T_j z_n)$. By applying Exercise 5.4.14 we get that

$$T_j^{(n)}(x_1, \ldots, x_n) \xrightarrow{\text{weak}} T^{(n)}(x_1, \ldots, x_n).$$

By Theorem 7.1.16 there exist convex combinations $T'_k$ of the $T_j$ such that

$$T'_k(x_1, \ldots, x_n) \to T(x_1, \ldots, x_n).$$

As this can be done for any finite $F \subset \mathcal{X}$, we can get a net $\{T'_{F,\varepsilon}\}$ such that $T'_{F,\varepsilon} \to T$ pointwise in norm. □

**9.1.3. Exercises.**

(9.1.1) Show that if $T, S \in \mathcal{B}(\mathcal{X})$ and both $TS$ and $ST$ are invertible, then both $T$ and $S$ are invertible. Find an example where $TS$ is invertible, but neither $T$ nor $S$ are.

(9.1.2) Show that if $\dim \mathcal{X} < \infty$, then any linear $T : \mathcal{X} \to \mathcal{X}$ is bounded.

(9.1.3) Show that if $\dim \operatorname{ran} T < \infty$, it is not necessarily true that $T$ is bounded.

(9.1.4) Let $\mathcal{X}$ be a normed space and $A \in \mathcal{B}(\mathcal{X})$ invertible. Show that

$$\|A\|\,\|A^{-1}\| = \sup\left\{\frac{\|Ax\|}{\|Ay\|} :\ x, y \in \mathcal{X},\ \|x\| = \|y\|\right\}.$$

(9.1.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{X} \to \mathcal{Y}$ linear, injective, and such that $T\big(\overline{B_1^{\mathcal{X}}(0)}\big) = \overline{B_1^{\mathcal{Y}}(0)}$. Show that $T$ is isometric.

(9.1.6) *(this exercise does not require sophisticated ideas but it is far from trivial; we include it here because it is where it corresponds topic-wise, but it likely requires more experience with operators)*

Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$. We define numbers

$$a(T) = \min\{k \in \mathbb{N} :\ \ker T^k = \ker T^{k+1}\}$$

and

$$d(T) = \min\{k \in \mathbb{N} :\ \operatorname{ran} T^k = \operatorname{ran} T^{k+1}\}.$$

(i) Show that if both $a(T)$ and $d(T)$ are finite, then they are equal.

(ii) If $n = a(T) = d(T)$, show that $\operatorname{ran} T^n$ is closed and that $\mathcal{X} = \operatorname{ran} T^n \oplus \ker T^n$.

(9.1.7) Show that (9.2) does indeed correspond to the operator-block matrix form of $ST$.

(9.1.8) Prove Proposition 9.1.3.

## 9.2. Banach Algebras and the Spectrum

While in this text we focus on complex vector spaces, the big majority of results hold for real spaces too. The spectrum, though, is a part of the story where differences are significant.

Even though the spectrum is mostly used to discuss operators, the definitions and main results work on any Banach algebra. Hence we consider Banach algebras in this section, as they provide a more general setting where we can still define and analyze the spectrum. Note that $\mathcal{B}(\mathcal{X})$ is a Banach algebra.

**9.2.1. Definition.**

*A **Banach Algebra** is a Banach space $\mathcal{A}$ that has this additional structure:*

- *a multiplication that makes it an algebra (that is, $\mathcal{A}$ is a ring where the multiplication is compatible with the vector space structure, in the sense that $\lambda(ab) = (\lambda a)b = a(\lambda b)$ for all $\lambda \in \mathbb{C}$ and $a, b \in \mathcal{A}$).*
- *for all $a, b \in \mathcal{A}$, we have $\|ab\| \leq \|a\| \, \|b\|$.*

If $\mathcal{A}$ satisfies all the conditions in Definition 9.2.1 with the exception of being complete, we say that $\mathcal{A}$ is a **normed algebra**.

An algebra $\mathcal{A}$ is **unital** if it has a nonzero element (necessarily unique) $I_{\mathcal{A}}$ such that $aI_{\mathcal{A}} = I_{\mathcal{A}}a$ for all $a \in \mathcal{A}$. In cases where it is clear by context, we may write $I$ instead of $I_{\mathcal{A}}$. When $\mathcal{A}$ is unital, we say that $a \in \mathcal{A}$ is **invertible** if there exists $b \in \mathcal{A}$ with $ab = ba = I_{\mathcal{A}}$; such $b$—when it exists—is unique and we denote it by $a^{-1}$. Motivated by the case where $\mathcal{A}$ is an algebra of operators, we write

$$GL(\mathcal{A}) = \{a \in \mathcal{A} : \ a \text{ is invertible}\}.$$

$\mathcal{A}$ is **abelian** (or **commutative**) if $ab = ba$ for all $a, b \in \mathcal{A}$.

The arguments used in Proposition 6.2.3 apply to any unital Banach algebra. So we have

***9.2.2. Proposition.***

*Let $\mathcal{A}$ be a unital Banach algebra. The set $GL(\mathcal{A})$ is open. More concretely, if $a \in GL(\mathcal{A})$ and $\|a - b\| \leq \|a^{-1}\|^{-1}$, then $b \in GL(\mathcal{A})$.*

*Proof.* Exercise 9.2.9. □

*9.2.3. Examples.* Let us see that we already know several examples of Banach algebras.

(a) For any locally compact Hausdorff space $T$, the algebra $C_0(T)$, with the supremum norm, is an abelian Banach algebra. If $T$ is compact, then $C_0(T) = C(T)$ is unital.

(b) For any locally compact Hausdorff space $T$, the algebra $C_b(T)$ of bounded continuous functions on $T$ is an abelian unital Banach algebra.

(c) Given any Banach space $\mathcal{X}$, the algebra $\mathcal{B}(\mathcal{X})$ of bounded operators on $\mathcal{X}$, with the operator norm, is a Banach algebra. Particular cases of this are $\mathcal{B}(\mathcal{H})$ for a Hilbert space $\mathcal{H}$, and more particularly $M_n(\mathbb{C})$, which we identify as $\mathcal{B}(\mathbb{C}^n)$. Any subalgebra of $M_n(\mathbb{C})$ is then an example of a Banach algebra. Note that even though all norms are equivalent on $M_n(\mathbb{C})$, most will fail to be submultiplicative so they will not make it a Banach algebra; the operator norm is submultiplicate, though, and it is the one considered here.

(d) The space $\ell^\infty(\mathbb{N})$, with pointwise addition and multiplication and the supremum norm, is a unital abelian Banach algebra. More generally, the space $L^\infty(\Omega, \Sigma, \mu)$ is a unital abelian Banach algebra for any measure space $(\Omega, \Sigma, \mu)$.

(e) The **disk algebra** $\mathbb{A} = \{f \in C(\overline{\mathbb{D}}) : \ f \ \text{ is analytic on } \mathbb{D}\}$ is an abelian unital Banach algebra with the norm $\|f\|_\infty = \max\{|f(z)| : \ z \in \overline{\mathbb{D}}\}$.

(f) For $p \in [1, \infty]$ the space $\ell^p(\mathbb{N})$, with pointwise multiplication, is a Banach algebra. Indeed, when $p < \infty$

$$\|xy\|_p^p = \sum_k |x_k y_k|^p \leq \|x\|_\infty^p \, \|y\|_p^p \leq \|x\|_p^p \, \|y\|_p^p,$$

so $\|xy\|_p \leq \|x\|_p \|y\|_p$. And when $p = \infty$, $\|xy\|_\infty \leq \|x\|_\infty \|y\|_\infty$ directly.

In all examples above, when the algebra is unital we have that $\|I_\mathcal{A}\| = 1$. This does not have to be the case in general, though. For an artificial way to construct an example, note that if you multiply the norm of any Banach algebra by a fixed number greater than 1 (so that the submultiplicativity is preserved), you get an equivalent norm. As we always have $\|I_\mathcal{A}\| = \|I_\mathcal{A}^2\| \leq \|I_\mathcal{A}\|^2$, this forces $\|I_\mathcal{A}\| \geq 1$. We can always go back to the case where $\|I_\mathcal{A}\| = 1$, though, by replacing the original norm with an equivalent norm. Indeed,

***9.2.4. Proposition.***

*Let $\mathcal{A}$ be a unital Banach algebra. Then there exists a norm $n$ on $\mathcal{A}$, equivalent to the original norm, such that $(A, n)$ is a unital Banach algebra and $n(I_\mathcal{A}) = 1$.*

*Proof.* We consider $\mathcal{B}(\mathcal{A})$, the space of bounded linear operators on $\mathcal{A}$, with the operator norm. Let $\Gamma : \mathcal{A} \to \mathcal{B}(\mathcal{A})$ be the map $\Gamma(a) = L_a$, where $L_a$ is the operator of left multiplication by $a$, $L_a(b) = ab$. Note that

$$\|L_a b\| = \|ab\| \leq \|a\| \, \|b\|, \tag{9.3}$$

so $L_a$ is bounded with $\|L_a\| \leq \|a\|$. Since $L_{a+b} = L_a + L_b$ and $L_{ab} = L_a L_b$, the map $\Gamma$ is an algebra homomorphism. It is also injective, for $L_a = 0$ implies $0 = aI_\mathcal{A} = a$. And it is unital, for $L_{I_\mathcal{A}} = I$.

Now we define $n(a) = \|\Gamma(a)\|$; that is, $n(a) = \|L_a\|$. This is a norm, by the linearity and injectivity of $\Gamma$; and it is submultiplicative since $\Gamma$ is multiplicative and the original norm is submultiplicative. We have shown that $n(a) \leq \|a\|$.

$$n(ab) = \|L_{ab}\| = \|L_a L_b\| \leq \|L_a\| \, \|L_b\| = n(a)n(b)$$

since the operator norm is submultiplicative, so $(\mathcal{A}, n)$ is a Banach algebra. And $n(I_{\mathcal{A}}) = \|L_{I_{\mathcal{A}}}\| = \|I\| = 1$. Finally, it remains to check that $n$ is equivalent to the original norm. First, $n(a) \geq \|L_a(I_{\mathcal{A}}/\|I_{\mathcal{A}}\|)\| = \|a\|/\|I_{\mathcal{A}}\|$. Using this inequality and (9.3),

$$\frac{1}{\|I_{\mathcal{A}}\|} \|a\| \leq n(a) \leq \|a\|. \qquad \square$$

**9.2.5. Definition.**

*Let $\mathcal{A}$ be a unital algebra and $a \in \mathcal{A}$. The* ***spectrum*** *of $a$ is the set*

$$\sigma(a) = \{\lambda \in \mathbb{C} : \; a - \lambda I \notin GL(\mathcal{A})\}.$$

*The* ***resolvent*** *of $a$ is the complement of the spectrum, denoted by $\rho(a)$. That is,*

$$\rho(a) = \{\lambda \in \mathbb{C} : \; a - \lambda I \in GL(\mathcal{A})\}.$$

*The* ***spectral radius*** *of $a$ is the number*

$$\operatorname{spr}(a) = \max\{|\lambda| : \; \lambda \in \sigma(a)\}.$$

An immediate consequence of Proposition 9.2.2 is that $\rho(a)$ is always open (Exercise 9.2.1), which means that $\sigma(a)$ is always closed. We will prove, after a few preliminaries, that $\sigma(a)$ is always nonempty and compact (Theorem 9.2.12).

*9.2.6. Remark*. An observation that we will use often is that if $\mu \in \mathbb{C}$, then $\sigma(T + \mu I) = \sigma(T) + \mu$. This is simply due to the fact that $(T + \mu I) - \lambda I = T - (\lambda - \mu)I$. So $\lambda \in \sigma(T + \mu I)$ if and only if $\lambda - \mu \in \sigma(T)$. Similarly, $\sigma(\mu T) = \mu \, \sigma(T)$.

*9.2.7. Examples*. There are some familiar cases where the spectrum is rather easy to determine—theoretically, at least.

(a) If $\mathcal{A} = M_n(\mathbb{C})$, and $a \in M_n(\mathbb{C})$, the spectrum $\sigma(a)$ consists of the eigenvalues of $a$.

(b) If $f \in L^\infty(X, \Sigma, \mu)$, then $\sigma(f) = \operatorname{ess\,ran} f$.

(c) If $T$ is compact Hausdorff and $\mathcal{A} = C(T)$, then for any $f \in \mathcal{A}$ we have

$$\sigma(f) = f(T).$$

(d) When $\mathcal{A}$ is not complete, the spectrum is often not very relevant. For instance let $\mathcal{A} = \mathbb{C}[x]$ with
$$\|p\| = \max\{|p(x)| : \ x \in [0,1]\}.$$
Then, as the only invertible polynomials are the constants, if $\deg p \geq 1$ then $\sigma(p) = \mathbb{C}$. Part of the problem here is that $\mathcal{A}$ is not Banach.

If one looks carefully at the definition of the spectrum of an element, it is apparent that $\sigma(a)$ depends on the algebra $\mathcal{A}$. Namely, if $a \in \mathcal{B} \subset \mathcal{A}$, it may be the case that for some $\lambda$ we have that $a - \lambda I \in GL(\mathcal{A})$, but $a - \lambda I \notin GL(\mathcal{B})$. A brutal example of this was given above: if for instance $\mathcal{A} = C[0,1]$ and $\mathcal{B} = \mathbb{C}[x]$, then $f(t) = 1 + t$ is invertible in $\mathcal{A}$ but not in $\mathcal{B}$. In this case $\sigma_{\mathcal{A}}(f) = [1,2]$, while $\sigma_{\mathcal{B}}(f) = \mathbb{C}$. This example is not entirely natural in the sense that $\mathcal{B}$ is not complete, but the problem still occurs with inclusions of Banach algebras. For instance, let $\mathcal{A} = C(\mathbb{T})$, and $\mathcal{B} \subset \mathcal{A}$ the disk algebra (cfr. Examples 9.2.3). If $f$ is the identity function $f(t) = t$, then
$$\sigma_{\mathcal{A}}(f) = \mathbb{T} \subsetneq \overline{\mathbb{D}} = \sigma_{\mathcal{B}}(f).$$
The dependence of the spectrum on the algebra is not totally random, though:

***9.2.8. Proposition.***

*Let $\mathcal{B} \subset \mathcal{A}$ be unital Banach algebras, with $I_{\mathcal{B}} = I_{\mathcal{A}}$. Let $b \in \mathcal{B}$. Then*

*(i) $GL(\mathcal{B})$ is relatively closed in $\mathcal{B} \cap GL(\mathcal{A})$;*

*(ii) $\sigma_{\mathcal{A}}(b) \subset \sigma_{\mathcal{B}}(b)$, and $\partial\sigma_{\mathcal{B}}(b) \subset \partial\sigma_{\mathcal{A}}(b)$;*

*(iii) if $\sigma_{\mathcal{A}}(b)$ has no holes, then $\sigma_{\mathcal{A}}(b) = \sigma_{\mathcal{B}}(b)$.*

*Proof.* Suppose that $\{b_n\} \subset GL(\mathcal{B})$ and $b_n \to b$ for some $b$ in $\mathcal{B} \cap GL(\mathcal{A})$. As $b \in GL(\mathcal{A})$, we have $b^{-1}b_n \to I$. Let $c_n = b^{-1}b_n$. Then $c_n \in GL(\mathcal{A})$ for $n$ big enough, and
$$\begin{aligned}\|I_{\mathcal{A}} - c_n^{-1}\| &\leq \|c_n(I_{\mathcal{A}} - c_n^{-1})\| + \|I_{\mathcal{A}} - c_n^{-1} - c_n(I_{\mathcal{A}} - c_n^{-1})\| \\ &= \|I_{\mathcal{A}} - c_n\| + \|(I_{\mathcal{A}} - c_n)(I_{\mathcal{A}} - c_n^{-1})\| \\ &\leq \|I_{\mathcal{A}} - c_n\| + \|I_{\mathcal{A}} - c_n\|\,\|I_{\mathcal{A}} - c_n^{-1}\|,\end{aligned}$$
Thus
$$\|I - c_n^{-1}\| \leq \frac{\|I - c_n\|}{1 - \|I - c_n\|}.$$
If we take $n$ big enough so that $\|I - c_n\| \leq 1/2$, then $\|I - c_n^{-1}\| \leq 2\|I - c_n\|$. Thus $b_n^{-1}b = c_n^{-1} \to I$, and so $b_n^{-1} \to b^{-1}$. This shows that $b \in GL(\mathcal{B})$ and so $GL(\mathcal{B})$ is relatively closed in $\mathcal{B} \cap GL(\mathcal{A})$.

From $GL(\mathcal{B}) \subset GL(\mathcal{A})$, we get that $\mathbb{C} \setminus \sigma_{\mathcal{B}}(b) \subset \mathbb{C} \setminus \sigma_{\mathcal{A}}(b)$; so $\sigma_{\mathcal{A}}(b) \subset \sigma_{\mathcal{B}}(b)$. And if $\lambda \in \partial\sigma_{\mathcal{B}}(b)$, then there exists $\{\lambda_n\} \subset \mathbb{C} \setminus \sigma_{\mathcal{B}}(b)$ with $\lambda_n \to \lambda$. Clearly $b - \lambda_n I \to b - \lambda I$; as $b - \lambda_n I \in GL(\mathcal{B})$ and $b - \lambda I \notin GL(\mathcal{B})$, by the first part of the proof we get that $b - \lambda I \notin GL(\mathcal{A})$. So $\lambda \in \sigma_{\mathcal{A}}(b)$ and, since $\sigma_{\mathcal{A}}(b) \subset \sigma_{\mathcal{B}}(b)$, we get that $\lambda \in \partial\sigma_{\mathcal{A}}(b)$. Thus $\partial\sigma_{\mathcal{B}}(b) \subset \partial\sigma_{\mathcal{A}}(b)$.

Finally, if $\sigma_{\mathcal{A}}(b)$ has no holes, then $\mathbb{C} \setminus \sigma_{\mathcal{A}}(b)$ is connected. As $GL(\mathcal{B})$ is relatively closed in $\mathcal{B} \cap GL(\mathcal{A})$, if $\lambda_n \to \lambda$ with $\lambda_n \in \mathbb{C} \setminus \sigma_{\mathcal{B}}(b)$ and $\lambda \in \mathbb{C} \setminus \sigma_{\mathcal{A}}(b)$, each $b - \lambda_n I \in GL(\mathcal{B})$ and $b - \lambda I = \lim_n b - \lambda_n I$, so $b - \lambda I \in GL(\mathcal{B})$, as it is clopen in $B \cap GL(\mathcal{A})$. Then $\lambda \in \mathbb{C} \setminus \sigma_{\mathcal{B}}(b)$, and $\mathbb{C} \setminus \sigma_{\mathcal{B}}(b)$ is clopen; as $\mathbb{C} \setminus \sigma_{\mathcal{A}}(b)$ is connected, we obtain $\mathbb{C} \setminus \sigma_{\mathcal{B}}(b) = \mathbb{C} \setminus \sigma_{\mathcal{A}}(b)$. So $\sigma_{\mathcal{B}}(b) = \sigma_{\mathcal{A}}(b)$. □

When $\mathcal{A}$ is an algebra, $a \in \mathcal{A}$, and $p \in \mathbb{C}[x]$, the expression $p(a)$ makes sense, and any algebraic manipulation that "works for $x$" also works for $a$. We show below that one can easily obtain the spectrum of $p(a)$ if the spectrum of $a$ is known.

**9.2.9. Proposition.** *(Spectral Mapping, Polynomial Version)*

*Let $\mathcal{A}$ be a unital algebra. If $a \in \mathcal{A}$ with $\sigma(a) \neq \varnothing$ and $p \in \mathbb{C}[x]$, then*

$$\sigma(p(a)) = p(\sigma(a)).$$

*If $a \in GL(\mathcal{A})$, then $\sigma(a^{-1}) = \{\lambda^{-1} : \lambda \in \sigma(a)\}$.*

*Proof.* If $p$ is constant, the equality is trivial. Otherwise, fix $\lambda \in \mathbb{C}$; we may then write, via the Fundamental Theorem of Algebra,

$$p(z) - \lambda = \mu_0(z - \mu_1) \cdots (z - \mu_n).$$

The coefficients $\mu_1, \ldots, \mu_n$ depend on $\lambda$. Since $\mathcal{A}$ is an algebra,

$$p(a) - \lambda I = \mu_0(a - \mu_1 I) \cdots (a - \mu_n I).$$

Note that $(a - \mu_j I)(a - \mu_k I) = (a - \mu_k I)(a - \mu_j I)$ for any $k, j$, so $p(a) - \lambda I \in GL(\mathcal{A})$ if and only if $\mu_1, \ldots, \mu_n \in \rho(a)$ (Exercise 9.2.8). Rephrasing, $p(a) - \lambda I \notin GL(\mathcal{A})$ if and only if $\mu_k \in \sigma(a)$ for some $k$. That is, $\lambda \in \sigma(p(a))$ if and only if $\mu_k \in \sigma(a)$ for some $k$; but $\lambda = p(\mu_k)$, so we have shown that $\lambda \in \sigma(p(a))$ if and only if $\lambda = p(\mu)$ for some $\mu \in \sigma(a)$.

When $a$ is invertible we have that $0 \notin \sigma(a) \cup \sigma(a^{-1})$. This allows us to consider $\lambda \in \mathbb{C} \setminus \{0\}$. Then

$$a - \lambda I = -\lambda a \left(a^{-1} - \frac{1}{\lambda}\right).$$

This shows that $\lambda \in \sigma(a)$ if and only if $\lambda^{-1} \in \sigma(a^{-1})$. □

*9.2.10. Remark* . If $a \in \mathcal{A}$ is **nilpotent**, that is if there exists $n \in \mathbb{N}$ with $a^n = 0$, then by Proposition 9.2.9 we get that $\sigma(a) = \{0\}$. When $\sigma(a) = \{0\}$, we say that $a$ is **quasinilpotent**.

We will need some basic but not-entirely-obvious estimations.

***9.2.11. Lemma.***

*Let $\mathcal{A}$ be a unital Banach algebra and $b \in GL(\mathcal{A})$. Then, for $h \in \mathcal{A}$ with $\|h\|$ small enough,*

$$\|(b+h)^{-1} - b^{-1}\| \le \|h\| \, \frac{\|b^{-1}\|^2}{1 - \|h\| \, \|b^{-1}\|}.$$

*In particular,* $\lim_{h \to 0}(b+h)^{-1} = b^{-1}$.

*Proof.* If $\|h\| < \|b^{-1}\|^{-1}$, then $\|(b+h) - b\| = \|h\| < \|b^{-1}\|^{-1}$, so $b + h \in GL(\mathcal{A})$ by Proposition 9.2.2.

If $\|a\| < 1$, then $I + a \in GL(\mathcal{A})$ and

$$\begin{aligned} 1 &= \|(I+a)(I+a)^{-1}\| = \|(I+a)^{-1} + a(I+a)^{-1}\| \\ &\ge \|(I+a)^{-1}\| - \|a(I+a)^{-1}\| \ge \|(I+a)^{-1}\| - \|a\| \, \|(I+a)^{-1}\|. \end{aligned}$$

Then

$$\|(I+a)^{-1}\| \le \frac{1}{1 - \|a\|}. \tag{9.4}$$

Now, using the equality $x^{-1} - y^{-1} = x^{-1}(y-x)y^{-1}$ and the inequality (9.4),

$$\begin{aligned} \|(b+h)^{-1} - b^{-1}\| &= \|(b+h)^{-1}(-h)b^{-1}\| \le \|h\| \, \|b^{-1}\| \, \|(b+h)^{-1}\| \\ &= \|h\| \, \|b^{-1}\| \, \|b^{-1}(I + hb^{-1})^{-1}\| \\ &\le \|h\| \, \|b^{-1}\|^2 \, \|(I + hb^{-1})^{-1}\| \\ &\le \|h\| \, \frac{\|b^{-1}\|^2}{1 - \|hb^{-1}\|} \le \|h\| \, \frac{\|b^{-1}\|^2}{1 - \|h\| \, \|b^{-1}\|}. \end{aligned}$$

□

The purpose of the above estimations is to be used in the crucial next result.

***9.2.12. Theorem.***

*Let $\mathcal{A}$ be a unital Banach algebra and $a \in \mathcal{A}$. Then $\sigma(a)$ is a nonempty compact set, and* $\operatorname{spr}(a) \le \|a\|$.

*Proof.* If $|\lambda| > \|a\|$, let $b = a/\lambda$; we have $\|b\| < 1$. By Lemma 6.2.1, $I - b$ is invertible, so $a - \lambda I = -\lambda(I - b)$ is also invertible. This shows that if $\lambda \in \sigma(a)$, then $|\lambda| \leq \|a\|$. That is, $\operatorname{spr}(a) \leq \|a\|$.

Recall that $\sigma(a)$ is always closed. Being bounded (by the first paragraph) and closed in $\mathbb{C}$, the spectrum $\sigma(a) \subset \mathbb{C}$ is always compact.

It remains to show that $\sigma(a) \neq \varnothing$. Fix $a \in \mathcal{A}$, $\varphi \in \mathcal{A}^*$ and define $f_{a,\varphi} : \rho(a) \to \mathbb{C}$ by

$$f_{a,\varphi}(z) = \varphi((zI - a)^{-1}).$$

If we consider Newton quotients for $f_{a,\varphi}$ (which make sense because $\rho(a)$ is open), we have for $z \in \rho(a)$ (and using the formula $a^{-1} - b^{-1} = a^{-1}(b-a)b^{-1}$, Lemma 9.2.11, and the continuity of $\varphi$)

$$\begin{aligned} \frac{f_{a,\varphi}(z+h) - f_{a,\varphi}(z)}{h} &= \varphi\left(\frac{((z+h)I - a)^{-1} - (zI - a)^{-1}}{h}\right) \\ &= \varphi\left(\frac{((z+h)I - a)^{-1}(-hI)(zI - a)^{-1}}{h}\right) \\ &= \varphi\left(-((z+h)I - a)^{-1}(zI - a)^{-1}\right) \\ &\xrightarrow[h \to 0]{} \varphi\left(-(zI - a)^{-2}\right). \end{aligned}$$

This shows that the complex derivative exists in all of $\rho(a)$ and so $f_{a,\varphi}$ is holomorphic on $\rho(a)$. Using (9.4) twice we have

$$|f_{a,\varphi}(z)| \leq \|\varphi\| \frac{1}{(|z| - \|a\|)^2},$$

for $|z| > \|a\|$, so $\lim_{|z| \to \infty} f_{a,\varphi}(z) = 0$, and thus $f_{a,\varphi}$ is bounded. If $\sigma(a) = \varnothing$, then $\rho(a) = \mathbb{C}$ and $f_{a,\varphi}$ is a bounded entire function; by Liouville's Theorem (Corollary 3.7.4), $f_{a,\varphi} = 0$. This we can do for all $\varphi \in \mathcal{A}^*$, and so for $|z| > \|a\|$—where we have already proven that $zI - a$ is invertible—we would obtain that $(zI - a)^{-1} = 0$, a contradiction. The conclusion is that $\sigma(a) \neq \varnothing$. ☐

*9.2.13. Remark*. The proof of the fact that $\sigma(a) \neq \varnothing$ for each $a \in \mathcal{A}$ is one significant case where the use of complex numbers is essential. Indeed, for real Banach algebras the spectrum can be empty even in finite dimension. Possibly the simplest example of a Banach algebra with an element with empty spectrum can be achieved by considering

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in M_2(\mathbb{R}).$$

Embedding this example in bigger spaces leads to examples of the same phenomenon in any dimension $\geq 2$.

Here is a straightforward application of Theorem 9.2.12:

***9.2.14. Proposition. (Gelfand–Mazur)***

*Let $\mathcal{A}$ be a unital Banach algebra such that every nonzero element is invertible. Then $\mathcal{A} = \mathbb{C}\,1$.*

*Proof.* Fix $a \in \mathcal{A}$. By Theorem 9.2.12 there exist $\lambda \in \mathbb{C}$ such that $a - \lambda I$ is not invertible. Hence $a - \lambda I = 0$, and $a = \lambda I$. $\square$

Given a Banach algebra $\mathcal{A}$ and $a, b \in \mathcal{A}$, it is not true in general that $\sigma(ab) = \sigma(ba)$. A canonical example of this is given by the unilateral shifts, mentioned briefly on Section 6.2 and considered in Example 9.5.10. The next best thing is true, though:

***9.2.15. Proposition.***

*Let $\mathcal{A}$ be a unital algebra and $a, b \in \mathcal{A}$. Then $\sigma(ab) \setminus \{0\} = \sigma(ba) \setminus \{0\}$.*

*Proof.* We will prove that $ab - \lambda I$ is invertible if and only if $ba - \lambda I$ is invertible, whenever $\lambda \neq 0$. We initially consider the case $\lambda = 1$. If $I - ab$ is invertible and $c = (I - ab)^{-1}$, then

$$(I + bca)(I - ba) = I + bca - ba - bcaba = I + bc(I - ab)a - ba = I + ba - ba = I.$$

A similar computation shows that $(I - ba)(I + bca) = I$. For arbitrary $\lambda \neq 0$, $ab - \lambda I = -\lambda(I - a(b/\lambda))$ and we can apply the previous case. The reverse implication is obtained by exchanging the roles of $a$ and $b$. $\square$

***9.2.16. Theorem.***

*Let $\mathcal{A}$ be a unital Banach algebra, and $a \in \mathcal{A}$. Then*

$$\operatorname{spr}(a) = \lim_{n \to \infty} \|a^n\|^{1/n}.$$

*Proof.* Given $\lambda \in \sigma(a)$, we know by Proposition 9.2.9 that $\lambda^n \in \sigma(a^n)$. By Theorem 9.2.12, $|\lambda^n| \leq \|a^n\|$. Then $|\lambda| \leq \|a^n\|^{1/n}$ and so $\operatorname{spr}(a) \leq \|a^n\|^{1/n}$. Hence

$$\operatorname{spr}(a) \leq \liminf_{n \to \infty} \|a^n\|^{1/n}. \tag{9.5}$$

Again we fix $\varphi \in \mathcal{A}^*$ and we consider $f_{a,\varphi}(z)$ as defined in the proof of Theorem 9.2.12. For $z \in \rho(a)$ this function is holomorphic and goes to zero at infinity; so the function $g_{a,\varphi}(z) = f_{a,\varphi}(1/z)$ is holomorphic for every $z$ with $z^{-1} \in \rho(a)$; in particular it is holomorphic on the disk of radius $\|a\|^{-1}$ around 0, where it admits a power series representation $g_{a,\varphi}(z) = \sum_{n=0}^{\infty} t_n z^n$. We can actually calculate this power series explicitly. Indeed, if $|z| < \|a\|^{-1}$, then

$\|za\| < 1$ and so the series $\sum_{k=0}^{\infty} z^k a^k$ converges and is the inverse of $I - za$. So

$$\left(\frac{1}{z}I - a\right)^{-1} = z(I - za)^{-1} = \sum_{k=0}^{\infty} z^{k+1}a^k = \sum_{k=1}^{\infty} a^{k-1}z^k.$$

Then, since $\varphi$ is bounded,

$$g_{a,\varphi}(z) = \varphi\Big(\left(\frac{1}{z}I - a\right)^{-1}\Big) = \sum_{k=1}^{\infty} \varphi(a^{k-1})z^k, \tag{9.6}$$

showing that $t_0 = 0$ and $t_n = \varphi(a^{n-1})$ for $n \geq 1$.

The function $g_{a,\varphi}(z)$ is analytic on $\{0\} \cup V$, where $V = \{z : z^{-1} \in \rho(a)\}$. The set $\{0\} \cup V$ contains the open disk with radius $\operatorname{dist}(0, V^c)$; indeed, if $|z| < |w|$ for all $w \in V^c$ then either $z = 0$ or $z \in V$. Hence the radius of convergence $r$ of the series (9.6) is at least

$$r \geq \operatorname{dist}(0, V^c) = \operatorname{dist}\left(0, \left\{\frac{1}{z} : z \in \sigma(a)\right\}\right) = \inf\left\{\frac{1}{|z|} : z \in \sigma(a)\right\}$$
$$= \frac{1}{\sup\{|z| : z \in \sigma(a)\}} = \frac{1}{\operatorname{spr}(a)}.$$

So, for any $z$ with $|z| < 1/\operatorname{spr}(a)$, the series in (9.6) converges. For such $z$, $\varphi(a^k)z^k \to 0$, and thus $\sup\{|\varphi(a^k z^k)| : k \in \mathbb{N}\} < \infty$ for all $\varphi \in A^*$. By Theorem 6.3.16 (Uniform Boundedness Principle) applied to the maps $\varphi \longmapsto \varphi(a^n z^n)$, we obtain that $R = \sup\{\|a^n z^n\| : n \in \mathbb{N}\} < \infty$. Then $\|a^n\|^{1/n} \leq R^{1/n}/|z|$, and

$$\limsup \|a^n\|^{1/n} \leq \frac{1}{|z|}.$$

As this inequality holds for any $z$ with $1/|z| > \operatorname{spr}(a)$, we have shown that

$$\limsup \|a^n\|^{1/n} \leq \operatorname{spr}(a). \tag{9.7}$$

Putting (9.5) and (9.7) together,

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq \operatorname{spr}(a) \leq \liminf_{n\to\infty} \|a^n\|^{1/n}.$$

This implies that $\lim_n \|a^n\|^{1/n}$ exists and agrees with $\operatorname{spr}(a)$. □

*9.2.17. Remark* . The spectral radius can be very different from the norm. Indeed, for any nontrivial Banach space there exist operators $a$ with $\|a\| = 1$ and $\operatorname{spr}(a) = 0$ (Exercise 9.5.2). For instance, in $M_2(\mathbb{C})$, let

$$a = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

As already mentioned, such operators are called **quasinilpotent**.

*9.2.18. Remark*. In the same vein as Remark 9.2.17, it is possible to have a Banach algebra $\mathcal{A}$ and $a, b \in \mathcal{A}$ such that $\|a - b\|$ is arbitrarily small, while $\sigma(a) = \{0\}$ and $1 \in \sigma(b)$ (Exercise 9.2.16).

**9.2.19. Definition.**

*A subalgebra $J \subset \mathcal{A}$ is an **ideal** if it is a closed two-sided ideal. When we need to refer to a non-necessarily closed ideal, we will use the terminology **algebraic ideal**.*

*When $\mathcal{A}$ has no proper ideals, we say that it is **simple**.*

Given an ideal $J \subset \mathcal{A}$, we can consider the Banach space quotient $\mathcal{A}/J$; we showed that this is again a Banach space in Proposition 5.3.13. It turns out that $\mathcal{A}/J$ is also a Banach algebra. What we need to check is that it is an algebra, and that the norm is submultiplicative. For the former one checks as usual that the product of representatives is independent of the choice of representative. And for the norm, given that the infimum is multiplicative on sets of positive numbers,

$$\begin{aligned}
\|(a+J)(b+J)\| &= \|ab+J\| = \inf\{\|ab+j\| : \ j \in J\} \\
&\leq \inf\{\|ab+jb+aj'+jj'\| : \ j, j' \in J\} \\
&= \inf\{\|(a+j)(b+j')\| : \ j, j' \in J\} \\
&\leq \inf\{\|a+j\| \, \|a+j'\| : \ j, j' \in J\} \\
&= \inf\{\|a+j\| : \ j \in J\} \inf\{\|b+j'\| : \ j' \in J\} \\
&= \|a+J\| \, \|b+J\|.
\end{aligned}$$

Let us collect some information about maximal ideals.

***9.2.20. Lemma.***

*Let $\mathcal{A}$ be a unital Banach algebra. Then*

*(i) every proper ideal is contained in a proper maximal ideal;*

*(ii) maximal ideals are closed;*

*(iii) if $\mathcal{A}$ is abelian, an ideal $J \subset \mathcal{A}$ is maximal if and only if $\dim \mathcal{A}/J = 1$.*

*The converse of (iii) holds even if $\mathcal{A}$ is not abelian.*

*Proof.* Let $J \subset \mathcal{A}$ be a proper algebraic ideal. Let $\mathcal{J}$ be the family of increasing collections of proper ideals that contain $J$, ordered by inclusion. Given an increasing family of these collections—a chain—the union is again an increasing family of ideals that contain $J$, so it is an upper bound, in $\mathcal{J}$, for the chain. Thus Zorn's lemma applies: there exists a maximal collection of proper ideals that contain $J$, say $\{J_\alpha\}_\alpha$. Define $J_{\max} = \bigcup_\alpha J_\alpha$. Because the union is increasing, this is an ideal, and it contains $J$. And it has to be proper: if $J_{\max} = \mathcal{A}$, then $I_\mathcal{A} \in J_{\max}$, which in turn means that there exists $\alpha$ with $I_\mathcal{A} \in J_\alpha$, contradicting that $J_\alpha$ is proper. So $J_{\max}$ is proper.

Since $J$ is proper, every nonzero $a \in J$ is not invertible. By Proposition 9.2.2, we have that $\|I_\mathcal{A} - a\| \geq 1$. That is, $\operatorname{dist}(I_\mathcal{A}, J) = 1$. Then $I_\mathcal{A} \notin \overline{J}$, so $\overline{J}$ is proper. In particular, maximal ideals are closed.

If $\dim \mathcal{A}/J = 1$, let $J' \supset J$ be an ideal. Given $b \in J' \setminus J$, then $b + J$ is nonzero in $\mathcal{A}/J$, and so $b + J$ and $I_\mathcal{A} + J$ are linearly dependent; thus there exists nonzero $\lambda \in \mathbb{C}$ with $b + J = \lambda(I_\mathcal{A} + J)$. So $b \in \lambda I_\mathcal{A}$+J; there exists $j \in J$ with $b = \lambda I_\mathcal{A} + j$. This we can write as $I_\mathcal{A} = \frac{1}{\lambda}(b - J) \in J'$, showing that $J' = \mathcal{A}$. So $J$ is maximal.

Conversely, suppose that $\mathcal{A}$ is abelian and that $J \subset \mathcal{A}$ is a maximal ideal. Let $b + J \in \mathcal{A}/J$ be nonzero, that is $b \notin J$. Consider the ideal $(b + J)\mathcal{A}/J$ in $\mathcal{A}/J$. If this ideal is all of $\mathcal{A}/J$, then there exists $c \in \mathcal{A}$ with $(b + J)(c + J) = I_\mathcal{A} + J$. This shows that every nonzero element in $\mathcal{A}/J$ is invertible. By Proposition 9.2.14, $\mathcal{A}/J = \mathbb{C}(I_\mathcal{A} + J)$, so $\dim \mathcal{A}/J = 1$. ☐

When $\mathcal{A}$ is not abelian, the codimension of $J$ can be anything. For instance, if $\mathcal{A} = M_n(\mathbb{C})$, then the only ideal is $\{0\}$, with codimension $n^2$. When $\mathcal{A} = \mathcal{B}(\mathcal{H})$, the algebra of bounded operators on an infinite-dimensional separable Hilbert space, the algebra $\mathcal{K}(\mathcal{H})$ of compact operators is the only nonzero ideal, thus maximal; and $\dim \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H}) = \infty$—in fact, this quotient, called the Calkin Algebra is nonseparable, so it is uncountably-dimensional.

It happens in many cases that a Banach algebra $\mathcal{A}$ is not unital; typical examples are $C_0(T)$ for a non-compact locally compact Hausdorff space $T$, and $K(\mathcal{H})$, the compact operators on a Hilbert space $\mathcal{H}$. It turns out that it is always possible to enlarge the algebra so that it has a unit. This can be done in many ways, but we want it to be "tight"; that is, we want to enlarge the algebra the least possible. For instance, we can enlarge $C_0(T)$ by going to the Stone–Čech compactification $C(\beta T)$; this would be brutal, as this unital extension is much, much larger—for instance, even if $C_0(T)$ is separable, $C(\beta T)$ will not be separable.

The least we need in this extension is the unit: if we have elements $a + \lambda I$, their sums and products are of the same form, so we may not need to attach anything else. Note that if we want this to happen in an algebra, necessarily

$$(a + \lambda I)(b + \mu I) = ab + \lambda b + \mu a + \lambda\mu,$$

which tells us what the multiplication should be.

**9.2.21. Proposition.** *(Unitization)*

*Let $\mathcal{A}$ be a non-unital Banach algebra. Then there exists a unital Banach algebra $\tilde{\mathcal{A}}$, unique up to isomorphism, such that*

*(i) there exists an isometric monomorphism $\pi : \mathcal{A} \to \tilde{\mathcal{A}}$;*

*(ii) $\pi(\mathcal{A})$ is a maximal ideal in $\tilde{\mathcal{A}}$, of codimension 1;*

*(iii) if $\mathcal{A}$ is abelian, so is $\tilde{\mathcal{A}}$.*

*Proof.* Let $\tilde{\mathcal{A}} = \mathcal{A} \oplus \mathbb{C}$, with pointwise addition and multiplication by scalars, and the multiplication

$$(a, \lambda)(b, \mu) = (ab + \lambda b + \mu a, \lambda\mu).$$

Then $\tilde{\mathcal{A}}$ is a unital algebra with unit $(0, 1)$. Define a norm by

$$\|(a, \lambda)\| = \|a\| + |\lambda|.$$

It is clear that this is a norm; and the embedding $\pi : a \longmapsto (a, 0)$ is isometric.

The algebra $\tilde{\mathcal{A}}$ is also complete: if $\{(a_n, \lambda_n)\}$ is a Cauchy sequence in $\tilde{\mathcal{A}}$, then $\|a_n - a_m\| \leq \|(a_n, \lambda_n) - (a_m, \lambda_m)\|$ so that $\{a_n\}$ is Cauchy in $\mathcal{A}$. As $\mathcal{A}$ is complete, there exists $a \in \mathcal{A}$ with $a = \lim a_n$. Similarly, there exists $\lambda \in \mathbb{C}$ with $\lambda = \lim \lambda_n$, and so $(a, \lambda) = \lim_n (a_n, \lambda_n)$, and $\tilde{\mathcal{A}}$ is complete.

That $\tilde{\mathcal{A}}$ is abelian when $\mathcal{A}$ is abelian is straightforward, and also that $\pi(\mathcal{A})$ is an ideal. And $\pi(\mathcal{A})$ is a maximal ideal by Lemma 9.2.20, since $\tilde{\mathcal{A}}/\pi(\mathcal{A}) = \mathbb{C}$.

It remains to consider the uniqueness. Suppose that $\tilde{\mathcal{A}}$ and $\hat{\mathcal{A}}$ satisfy the above conditions, with isometric monomorphisms $\pi : \mathcal{A} \to \tilde{\mathcal{A}}$ and $\hat{\pi} : \mathcal{A} \to \hat{\mathcal{A}}$. By Lemma 9.2.20 the quotient $\tilde{\mathcal{A}}/\pi(\mathcal{A})$ is one dimensional. So for each $a \in \tilde{\mathcal{A}}$ there exists a unique scalar $\tilde{\tau}(a)$ such that $a + \pi(\mathcal{A}) = \tilde{\tau}(a) + \pi(\mathcal{A})$. It is straightforward to see that $\tilde{\tau}$ is linear and multiplicative, that $\tau(1) = 1$, and that $\pi(A) = \ker \tilde{\tau}$. Similarly, there exists a corresponding $\hat{\tau}$ for $\hat{\mathcal{A}}$. Then

$$\tilde{\mathcal{A}} = \{\lambda 1 + \pi(a) : \ \lambda \in \mathbb{C}, \ a \in \mathcal{A}\}$$

and

$$\hat{\mathcal{A}} = \{\lambda 1 + \hat{\pi}(a) : \ \lambda \in \mathbb{C}, \ a \in \mathcal{A}\}.$$

Let $\gamma : \tilde{\mathcal{A}} \to \hat{\mathcal{A}}$ be given by $\gamma(\lambda 1 + \pi(a)) = \lambda 1 + \hat{\pi}(a)$. This is injective and well-defined, for if $\lambda 1 + \pi(a) = \lambda' 1 + \pi(a')$, then $(\lambda - \lambda')1 = \pi(a - a')$; applying $\tau$ we get $\lambda - \lambda' = 0$, and this forces $\pi(a - a')$ which gives $a - a' = 0$ by the injectivity of $\pi$. It is also surjective by construction. The uniqueness also makes $\gamma$ linear and multiplicative, and so it is an isomorphism of algebras. The way it is defined makes $\gamma$ automatically isometric. ∎

Now we can define the spectrum in arbitrary Banach algebras. Namely,

**9.2.22. Definition.**

*Let $\mathcal{A}$ be a non-unital Banach algebra, and $a \in \mathcal{A}$. The* ***spectrum*** *of $a$ is*

$$\sigma_{\mathcal{A}}(a) := \sigma_{\tilde{\mathcal{A}}}(a).$$

*A* ***character*** *is a nonzero, linear, multiplicative mapping $\tau : \mathcal{A} \to \mathbb{C}$. We denote the set of characters of $\mathcal{A}$ by $\Sigma(\mathcal{A})$.*

We have already encountered the characters of a Banach algebra, for the concrete case $\mathcal{A} = C(T)$, in Definition 7.4.3 and Proposition 7.4.6; we also just used them in the proof of Proposition 9.2.21. For arbitrary $\mathcal{A}$, the character space $\Sigma(\mathcal{A})$ is of little interest. For example, $\Sigma(M_n(\mathbb{C})) = \varnothing$ (Exercise 9.2.13). Characters become useful, as we will see below, when $\mathcal{A}$ is abelian.

***9.2.23. Proposition.***

*Let $\mathcal{A}$ be a unital abelian Banach algebra and $a \in \mathcal{A}$. Then the map $\tau \longmapsto \ker\tau$ is a bijection between $\Sigma(\mathcal{A})$ and the set of maximal ideals of $\mathcal{A}$.*

*Proof.* Fix $\tau \in \Sigma(\mathcal{A})$. Then $\ker\tau$ is an ideal in $\mathcal{A}$. Since $\tau$ is nonzero and $\tau(I_{\mathcal{A}}) = \tau(I_{\mathcal{A}}^2) = \tau(I_{\mathcal{A}})^2$, we have $\tau(I_{\mathcal{A}}) = 1$. For any $b \in \mathcal{A}$, we have $b - \tau(b)\, I_{\mathcal{A}} \in \ker\tau$, so $\mathcal{A} = \mathbb{C}\, I + \ker\tau$. Thus $\dim \mathcal{A}/J = 1$ and $\ker\tau$ is maximal by Lemma 9.2.20.

Conversely, let $J \subset \mathcal{A}$ be a maximal ideal; as $\mathcal{A}$ is abelian, Lemma 9.2.20 implies that $\mathcal{A}/J = \mathbb{C}(I_{\mathcal{A}} + J)$. Given any $a \in \mathcal{A}$, we define $\tau(a) \in \mathbb{C}$ as the number such that $a + J = \tau(a)\, I_{\mathcal{A}} + J$. This number is unique because $I_{\mathcal{A}} \notin J$, so $\tau$ is well-defined. The uniqueness implies that $\tau$ is a character (Exercise 9.2.10). ☐

An immediate consequence of Proposition 9.2.23 is, using Proposition 5.5.9, that characters are always bounded. They also allow us to characterize the spectrum:

***9.2.24. Proposition.***

*Let $\mathcal{A}$ be an abelian Banach algebra and $a \in \mathcal{A}$.*

*(i) If $\mathcal{A}$ is unital, then $\sigma(a) = \{\tau(a) : \ \tau \in \Sigma(\mathcal{A})\}$.*

*(ii) If $\mathcal{A}$ is non-unital, then $\sigma(a) = \{\tau(a) : \ \tau \in \Sigma(\mathcal{A})\} \cup \{0\}$.*

*Proof.* Assume first that $\mathcal{A}$ is unital. If $\tau \in \Sigma(\mathcal{A})$, then $\tau(a - \tau(a)\, I_{\mathcal{A}}) = 0$, so $a - \tau(a)\, I_{\mathcal{A}} \in \ker\tau$, a maximal ideal; thus $a - \tau(a)\, I_{\mathcal{A}}$ is not invertible, and $\tau(a) \in \sigma(a)$. Conversely, if $\lambda \in \sigma(a)$, then $a - \lambda\, I_{\mathcal{A}}$ is not invertible; by Lemma 9.2.20 there exists a proper maximal ideal $J$ with $a - \lambda I_{\mathcal{A}} \in J$ (just take the proper ideal $\mathcal{A}(a - \lambda I_{\mathcal{A}})$ and take a maximal ideal that contains it). By Proposition 9.2.23 there exists $\tau \in \Sigma(\mathcal{A})$ with $J = \ker\tau$. Then $0 = \tau(a - \lambda I) = \tau(a) - \lambda$, and $\lambda = \tau(a)$.

When $\mathcal{A}$ is not unital, we define the spectrum in the unitization $\tilde{\mathcal{A}}$. The result follows directly from the unital case and the following claim:

$$\Sigma(\tilde{\mathcal{A}}) = \{(a,\lambda) \longmapsto \tau(a) + \lambda : \tau \in \Sigma(\mathcal{A}) \cup \{0\}\}. \tag{9.8}$$

To verify (9.8), let $\eta \in \Sigma(\tilde{\mathcal{A}})$. Define $\tau(a) = \eta(a,0)$; then $\tau$ is a character on $\mathcal{A}$; then $\eta(a,\lambda) = \eta(a,0) + \eta(0,\lambda) = \tau(a) + \lambda$. Conversely, we need to check that $\tilde{\tau}(a,\lambda) = \tau(a) + \lambda$ is a character when $\tau$ is a homomorphism (even if zero). It is clearly linear, and we have

$$\begin{aligned}\tilde{\tau}((a,\lambda)(b,\lambda')) &= \tilde{\tau}(ab + \lambda b + \lambda' a, \lambda\lambda') = \tau(ab + \lambda b + \lambda' a) + \lambda\lambda' \\ &= \tau(a)\tau(b)\lambda\tau(b) + \lambda'\tau(a) + \lambda\lambda' = (\tau(a) + \lambda)(\tau(b) + \lambda') \\ &= \tilde{\tau}(a,\lambda)\tilde{\tau}(b,\lambda').\end{aligned}$$

□

### 9.2.1. Exercises.

(9.2.1) Prove that for every $a \in \mathcal{A}$, the resolvent $\rho(a)$ is open.

(9.2.2) If $\mathcal{A} = M_n(\mathbb{C})$ and $a \in M_n(\mathbb{C})$, show that the spectrum $\sigma(a)$ consists of the eigenvalues of $a$.

(9.2.3) Let $a \in \ell^\infty(\mathbb{N})$. Show that $\sigma(a) = \overline{\{a(n) :\ n \in \mathbb{N}\}}$.

(9.2.4) Let $f \in L^\infty(X, \Sigma, \mu)$. Show that $\sigma(f) = \operatorname{ess\,ran} f$.

(9.2.5) If $T$ is compact Hausdorff and $\mathcal{A} = C(T)$, show that for any $f \in \mathcal{A}$ we have

$$\sigma(f) = f(T).$$

(9.2.6) Let $T$ be a locally compact Hausdorff space, and $\mathcal{A} = C_0(T)$. Show that the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$ is $C(T_\infty)$, the continuous functions on the one-point compactification of $T$.

(9.2.7) If $T$ is locally compact Hausdorff and $f \in C_0(T)$, show that

$$\sigma(f) = \overline{f(T)} = \{0\} \cup f(T).$$

Do this is two ways: directly, and by using characters.

(9.2.8) Let $R$ be a unital ring, and $a_1, \ldots, a_m \in R$. Show that if $a_k a_j = a_j a_k$ for all $j, k$ then $a_1 \cdots a_m$ is invertible if and only if each $a_j$ is invertible.

(9.2.9) Prove Proposition 9.2.2.

(9.2.10) Complete the proof of Proposition 9.2.23 by showing that if $J \subset \mathcal{A}$ is maximal and $\tau(a)$ is the unique scalar such that $a + J = \tau(a) I_{\mathcal{A}} + J$, then $\tau$ is a character.

(9.2.11) Use Lemma 9.2.20 and Proposition 7.4.13 to give an alternative proof of Proposition 7.4.6.

(9.2.12) Consider the Banach algebra $\mathcal{A} = \ell^\infty(\mathbb{N})$ and $K \subset \mathbb{C}$ compact. Show that there exists $a \in \mathcal{A}$ with $\sigma(a) = K$.

(9.2.13) Show that $\Sigma(M_n(\mathbb{C})) = \varnothing$.

(9.2.14) Let $\mathcal{A}, \mathcal{B}$ be unital Banach algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a unital homomorphism. Show that, for any $a \in \mathcal{A}$, $\sigma(\pi(a)) \subset \sigma(a)$.

(9.2.15) Consider the Banach algebra $\mathcal{A} = M_n(\mathbb{C})$. Fix $\varepsilon > 0$ and define

$$A_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad B_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \varepsilon & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Show that $\sigma(A_n) = \{0\}$, $\sigma(B_n) = \{\omega : \ \omega^n = \varepsilon\}$, and $\|A_n - B_n\| = \varepsilon$.

(9.2.16) Let $\varepsilon > 0$. Use Exercise 9.2.15 to construct operators

$$A, B \in \mathcal{B}\Big(\bigoplus_{n=2}^\infty \mathbb{C}^n\Big)$$

such that

$$\|A - B\| = \varepsilon, \qquad \sigma(A) = \{0\}, \qquad \sigma(B) = \mathbb{T}.$$

## 9.3. The Riesz Functional Calculus

Given any unital Banach algebra $\mathcal{A}$ and $a \in \mathcal{A}$, from the fact that $\mathcal{A}$ is an algebra it makes sense to evaluate $p(a)$ for any polynomial $p \in \mathbb{C}[x]$. Our algebra also has a topological structure, so it might make sense to calculate $f(a)$ for a function that is a limit of polynomials. This works at different levels depending on the structure we have available. On a Banach algebra, we will restrict ourselves to holomorphic functions; in C*-algebras one is able to make sense of $f(a)$ for continuous functions (at the expense of being more restrictive on $a$), and for von Neumann algebras we can allow $f$ to be any Borel function, again at the expense of being restrictive on $a$.

For a simple case, when $\mathcal{A} = M_n(\mathbb{C})$ and $A \in \mathcal{A}$, if $A$ is diagonalizable we may write $A = SDS^{-1}$, with $D$ diagonal. And then $p(A) = Sp(D)S^{-1}$ for any $p \in \mathbb{C}[x]$. Note that if $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $A$ (counting multiplicities), then $p(D)$ is simply the diagonal matrix where the $j, j$ entry is $p(\lambda_j)$; so it makes sense to define $f(A) = Sf(D)S^{-1}$, where $f(D)$ is the diagonal matrix with diagonal $\{f(\lambda_j)\}_j$. This theme of evaluating $f$ on the spectrum of $a$ will reappear every time we talk about functional calculus.

The departing idea here is Cauchy's Integral Formula (Theorem 3.9.1): for $f$ holomorphic on an open set $V$, a rectifiable Jordan curve $\gamma$ in $V$ (so $\gamma$ is closed, finite length, no intersections), and any $w \in V$ that is inside $\gamma$ (that is, its winding number is 1), then

$$f(w) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z-w}\, dz. \tag{9.9}$$

The observation is that it is possible to make sense of the integral above when $w$ is in a Banach algebra. Namely, an integral as above is a limit of Riemman sums; and to guarantee the convergence it is enough that the curve is well-behaved and that the argument is continuous. Whether the integrand takes numerical values or values in a Banach algebra $\mathcal{A}$ does not change the proof of existence of the integral (other than using the norm of $\mathcal{A}$ instead of the absolute value).

Let us recap some definitions from Chapter 3. A **curve** $\gamma$ is a continuous map $\gamma : [0,1] \to \mathbb{C}$. It is **closed** if $\gamma(1) = \gamma(0)$. It is **simple** if it has no intersections: $\gamma(t_1) = \gamma(t_2)$ imply that either $t_1 = t_2$, or $t_1, t_2 \in \{0,1\}$. It is **rectifiable** if its arc-length is finite. Recall (Section 3.4) that for any $w \in \mathbb{C}$ its **winding number** with respect to $\gamma$ is the index

$$\operatorname{Ind}_\gamma(w) = \frac{1}{2\pi i} \int_\gamma \frac{1}{z-w}\, dz.$$

We proved in Proposition 3.4.2 that $\operatorname{Ind}_\gamma(w) \in \mathbb{Z}$. We say that a simple curve $\gamma$ is **positively oriented** if $\operatorname{Ind}_\gamma(w) \in \{0,1\}$ for all $w \in \mathbb{C}$. We say that $w$ is **inside** $\gamma$ if $\operatorname{Ind}_\gamma(w) = 1$, and **outside** if $\operatorname{Ind}_\gamma(w) = 0$. We will write $K \subset \gamma$ is $z$ is inside $\gamma$ for all $z \in K$. Note that this is not set inclusion per se, rather its the inclusion of $K$ in the region surrounded by $\gamma$.

A family $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ of closed simple rectifiable curves is **positively oriented** if $\gamma_j(s) \neq \gamma_k(t)$ for all $k \neq j$, and all $s, t \in [0,1]$ (the curves don't intersect), and for any $w \notin \mathbb{C} \setminus \bigcup_{j=1}^n \gamma_j$

$$\operatorname{Ind}_\gamma(w) := \sum_{j=1}^n \operatorname{Ind}_{\gamma_j}(w) \in \{0,1\}.$$

The set $\{w \in \mathbb{C} : \operatorname{Ind}_\gamma(w) = 1\}$ is the **interior** of $\gamma$.

Given $a \in \mathcal{A}$ we will denote by $H(a)$ the set of functions that are holomorphic on an open set containing $\sigma(a)$. This set is an algebra: if $\alpha \in \mathbb{C}$, $f, g \in H(a)$, then $\lambda f + g, fg \in H(a)$. In each case we take the domain to be the intersection of the domains of $f$ and $g$.

As mentioned above, the integral

$$\frac{1}{2\pi i} \int_\gamma f(z)\, (zI_\mathcal{A} - a)^{-1}\, dz \tag{9.10}$$

makes sense for any $a \in \mathcal{A}$, as long as $(zI - a)^{-1}$ is continuous. Lemma 9.2.11 guarantees that $(zI-a)^{-1}$ is continuous at $z$ for all $z \in \rho(a)$, and the compactness of $\gamma$ then makes the function $z \longmapsto (zI - a)^{-1}$ uniformly continuous. So, as long as $\gamma$ does not touch $\sigma(a)$, the integral (9.10) is well-defined. And the definition does not actually depend on $\gamma$:

**_9.3.1. Proposition._**

*Let $a \in \mathcal{A}$ and $\gamma, \eta$ be two families of simple, rectifiable, positively oriented curves such that $\sigma(a)$ lies inside both $\gamma$ and $\eta$. If $V$ is open, $\gamma, \eta \subset V$, and $f : V \to \mathbb{C}$ is holomorphic, then*

$$\int_\gamma f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz = \int_\eta f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz.$$

*Proof.* Write $\gamma = \{\gamma_1, \ldots, \gamma_n\}$, $\eta = \{\eta_1, \ldots, \eta_m\}$, and $\eta_j^{-1}$ for the reverse curve of $\eta_j$, that is $\eta_j^{-1}$ but traversed in the opposite direction. Let $U = V \setminus \sigma(a)$. Then $\mathbb{C} \setminus U = (\mathbb{C} \setminus V) \cup \sigma(a)$. If $w \in \mathbb{C} \setminus V$, then (as $\gamma, \eta \subset V$)

$$\operatorname{Ind}_{\gamma \cup \eta^{-1}}(w) = \sum_{j=1}^n \operatorname{Ind}_{\gamma_j}(w) - \sum_{j=1}^m \operatorname{Ind}_{\eta_j}(w) = 0 - 0 = 0.$$

And if $w \in \sigma(a)$, then

$$\operatorname{Ind}_{\gamma \cup \eta^{-1}}(w) = \sum_{j=1}^n \operatorname{Ind}_{\gamma_j}(w) - \sum_{j=1}^m \operatorname{Ind}_{\eta_j}(w) = 1 - 1 = 0.$$

So the interior of $\gamma \cup \eta^{-1}$ is inside $U$. Given $\varphi \in \mathcal{A}^*$, we know from Theorem 9.2.12 that the function $z \longmapsto \varphi((zI - a)^{-1})$ is analytic on $U$; thus so is $f(z)\,\varphi((zI - a)^{-1})$. By Cauchy's Theorem (Theorem 3.9.1), and using that $\varphi$ is continuous,

$$\begin{aligned} 0 &= \int_{\gamma \cup \eta^{-1}} f(z)\,\varphi((zI_{\mathcal{A}} - a)^{-1})\,dz \\ &= \int_\gamma f(z)\,\varphi((zI_{\mathcal{A}} - a)^{-1})\,dz - \int_\eta f(z)\,\varphi((zI_{\mathcal{A}} - a)^{-1})\,dz \\ &= \varphi\left(\int_\gamma f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz - \int_\eta f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz\right). \end{aligned}$$

As $\varphi$ was arbitrary, we get (for instance by Corollary 5.7.7)

$$\int_\gamma f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz = \int_\eta f(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz.$$

□

Note that an analog of Proposition 9.3.1 was not needed in Section 3.5 as we had $f(a)$, while here we are trying to use the Cauchy integral to **define** $f(a)$.

We can now state and prove the holomorphic functional calculus in a Banach algebra.

***9.3.2. Theorem. (Holomorphic Functional Calculus)***

*Let $a \in \mathcal{A}$. Then the map $\Gamma : H(a) \to \mathcal{A}$ given by $\Gamma : f \longmapsto f(a)$, where*

$$f(a) = \frac{1}{2\pi i} \int_\gamma f(z)\,(zI_\mathcal{A} - a)^{-1}\,dz$$

*and $\sigma(a) \subset \gamma$, is well-defined (that is, the integral exists and does not depend on $\gamma$), and*

*(i) $\Gamma$ is a homomorphism;*

*(ii) if $f \in H(a)$, and $f(z) = \sum_{j=0}^\infty c_j z^j$, then $\Gamma(f) = \sum_{j=0}^\infty c_j a^j$;*

*(iii) if $\{f_n\} \subset H(a)$ and $f_n \to f$ uniformly on compact subsets inside an open set that contains $\sigma(a)$, then $f_n(a) \to f(a)$.*

*The morphism $\Gamma : H(a) \to \mathcal{A}$ satisfying (i), (ii), (iii) is unique.*

*Proof.* The uniqueness is automatic by the continuity and the fact that $\Gamma(p) = p(a)$ for all polynomials $p$.

By Proposition 9.3.1, $\Gamma(f) = f(a)$ is well-defined and does not depend on the curve. That $\Gamma$ is linear follows directly from the definition of integral. To see that $\Gamma$ is multiplicative, let $f, g \in H(a)$ with domains $V_1$ and $V_2$, and choose $\gamma, \eta$ positively oriented families of closed, rectifiable, simple curves such that $\sigma(a) \subset \gamma \subset V_1$, $\sigma(a) \subset \eta \subset V_2$, and such that $\gamma$ is inside $\eta$. We will need the well-known formula

$$c^{-1} - b^{-1} = b^{-1}(b - c)c^{-1}.$$

Then, noting that $\int_\gamma \frac{f(z)}{w-z}\,dz = 0$ for $w \in \eta$ (as $f$ is analytic on the interior of $\gamma$ and $w$ is outside of $\gamma$)

$$\begin{aligned}
f(a)g(a) &= \left(\frac{1}{2\pi i}\int_\gamma f(z)\,(zI_\mathcal{A} - a)^{-1}\,dz\right)\left(\frac{1}{2\pi i}\int_\eta g(z)\,(zI_\mathcal{A} - a)^{-1}\,dz\right) \\
&= \frac{1}{(2\pi i)^2}\int_\gamma\int_\eta f(z)g(w)\,(zI_\mathcal{A} - a)^{-1}(wI_\mathcal{A} - a)^{-1}\,dw\,dz \\
&= \frac{1}{(2\pi i)^2}\int_\gamma\int_\eta f(z)g(w)\,\frac{(wI_\mathcal{A} - a)^{-1} - (zI_\mathcal{A} - a)^{-1}}{z - w}\,dw\,dz \\
&= \frac{1}{(2\pi i)^2}\left[\int_\gamma f(z)\,(zI_\mathcal{A} - a)^{-1}\int_\eta \frac{g(w)}{w - z}\,dw\,dz\right.
\end{aligned}$$

$$
\begin{aligned}
&\qquad - \int_\eta g(w)\,(wI_{\mathcal{A}} - a)^{-1} \int_\gamma \frac{f(z)}{w-z}\,dz\,dw\Big] \\
&= \frac{1}{(2\pi i)^2} \int_\gamma f(z)\,(zI_{\mathcal{A}} - a)^{-1} \int_\eta \frac{g(w)}{w-z}\,dw\,dz \\
&= \frac{1}{2\pi i} \int_\gamma f(z)g(z)\,(zI_{\mathcal{A}} - a)^{-1}\,dz = (fg)(a).
\end{aligned}
$$

Let $p(x) = \sum_{j=0}^n c_j x^j$. Choose $\gamma$ to be the circle of radius $2\|a\|$ centered at 0. Then, using that $\Gamma$ is a homomorphism, Lemma 6.2.1, and that uniform convergence allows us to interchange limits,

$$
\begin{aligned}
\Gamma(p) &= \frac{1}{2\pi i} \int_\gamma \sum_{j=0}^n c_j z^j\,(zI_{\mathcal{A}} - a)^{-1}\,dz = \sum_{j=0}^n c_j \frac{1}{2\pi i} \int_\gamma z^j\,(zI_{\mathcal{A}} - a)^{-1}\,dz \\
&= \sum_{j=0}^n c_j \left( \frac{1}{2\pi i} \int_\gamma z\,(zI_{\mathcal{A}} - a)^{-1}\,dz \right)^j = \sum_{j=0}^n c_j \left( \frac{1}{2\pi i} \int_\gamma (I_{\mathcal{A}} - a/z)^{-1}\,dz \right)^j \\
&= \sum_{j=0}^n c_j \left( \frac{1}{2\pi i} \int_\gamma \sum_{k=0}^\infty \frac{a^k}{z^k}\,dz \right)^j = \sum_{j=0}^n c_j \left( \sum_{k=0}^\infty \frac{1}{2\pi i} \int_\gamma \frac{1}{z^k}\,dz\,a^k \right)^j \\
&= \sum_{j=0}^n c_j a^j = p(a).
\end{aligned}
$$

Now suppose that $f_n \to f$ uniformly on compact sets in $V$ with $V \supset \sigma(a)$. Choose $\gamma$ with $\sigma(a) \subset \gamma \subset V$. By the compactness of $\gamma$ and the continuity of $z \longmapsto (zI_{\mathcal{A}} - a)^{-1}$ (Lemma 9.2.11), there exists $c > 0$ with $\|(zI_{\mathcal{A}} - a)^{-1}\| \le c$. Then

$$
\begin{aligned}
\|f_n(a) - f(a)\| &= \left\| \int_{\gamma_j} (f_n(z) - f(z))\,(zI_{\mathcal{A}} - a)^{-1}\,dz \right\| \\
&\le c\,|\gamma|\,\sup\{|f_n(z) - f(z)\| = c\,|\gamma|\,\|f_n - f\| \to 0.
\end{aligned}
$$

Finally, when $f$ is analytic it is a uniform limit of polynomials, and hence

$$
f(a) = \lim_{n\to\infty} \sum_{j=0}^n c_j a^j = \sum_{j=0}^\infty c_j a^j. \qquad \square
$$

***9.3.3. Theorem. (Spectral Mapping, Riesz version)***

*Let $a \in \mathcal{A}$, $f \in H(a)$. Then $\sigma(f(a)) = f(\sigma(a))$.*

*Proof.* Let $\lambda \in \sigma(a)$. The function $f(z) - f(\lambda)$ is analytic with a zero at $z = \lambda$, so by Corollary 3.6.3 $f(z) - f(\lambda) = (z - \lambda)\,g(z)$ for some holomorphic $g$. Then, since the Riesz functional calculus is a homomorphism that maps the identity $z$ to $a$,

$$f(a) - f(\lambda)I = (a - \lambda I)\, g(a).$$

As $(a-\lambda I)$ commutes with $g(a)$ and it is not invertible, if follows that $f(a) - f(\lambda)I$ is not invertible. So $f(\lambda) \in \sigma(f(a))$ and we conclude that $f(\sigma(a)) \subset \sigma(f(a))$.

Conversely, if $\mu \notin f(\sigma(a))$, then $f(z)-\mu$ is analytic and $|f(z)-\mu| \geq \delta > 0$ for all $z$ on an open set containing $f(\sigma(a))$. So it is invertible, and then $f(a) - \mu I$ is invertible. Thus $\mu \notin \sigma(f(a))$. Then $\sigma(f(a)) \subset f(\sigma(a))$. $\square$

Here is an application of Riesz Functional Calculus that will be needed later. A canonical example of a Banach algebra is the algebra $\mathcal{B}(\mathcal{X})$ of bounded operators over a Banach space $\mathcal{X}$. We emphasize that the holomorphic function $f_\lambda$ used in Proposition 9.3.4 is trivial to construct, for if $\lambda$ is isolated in $\sigma(T)$ then we can take an open set $V$ with two connected components $V_1$ and $V_2$, with $\lambda \in V_1$ and $\sigma(T) \setminus \{\lambda\} \subset V_2$; then one simply defines $f_\lambda = 1$ on $V_1$ and $f_\lambda = 0$ on $V_2$.

**_9.3.4. Proposition. (Riesz Projections)_**

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$, and $\lambda \in \sigma(T)$ isolated. Then $\lambda$ is an eigenvalue for $T$, and if $U, V$ are disjoint open sets with $\lambda \in U$ and $\sigma(a) \setminus \{\lambda\} \subset V$, with $f_\lambda$ an holomorphic function on $U \cup V$ such that $f_\lambda|_U = 1$, $f_\lambda|_V = 0$, the operator*

$$P_\lambda = f_\lambda(T) = \frac{1}{2\pi i} \int_\gamma f_\lambda(z)\, (zI - T)^{-1}\, dz,$$

*is a nonzero idempotent with range the eigenspace $X_\lambda = \{x \in \mathcal{X} : \ Tx = \lambda x\}$. In particular, $TP_\lambda = \lambda P_\lambda$.*

*Proof.* We know that $P_\lambda \neq 0$ by Theorem 9.3.3, since $\sigma(P_\lambda) = f_\lambda(\sigma(T)) = \{0, 1\}$. And $P_\lambda$ is a projection, since $f_\lambda^2 = f_\lambda$; we have, since functional calculus is a homomorphism,

$$P_\lambda^2 = f_\lambda(T) f_\lambda(T) = f_\lambda^2(T) = f_\lambda(T) = P_\lambda.$$

As $z f_\lambda(z) = \lambda f_\lambda(z)$,

$$TP_\lambda = \frac{1}{2\pi i} \int_\gamma z\, f_\lambda(z)\, (zI - T)^{-1}\, dz = \frac{1}{2\pi i} \int_\gamma \lambda\, f_\lambda(z)\, (zI - T)^{-1}\, dz = \lambda P_\lambda.$$

Then any $x \in P_\lambda \mathcal{H}$ is an eigenvector for $\lambda$; indeed, if $P_\lambda x = x$, then $Tx = TP_\lambda x = \lambda P_\lambda x = \lambda x$. And conversely, if $Tx = \lambda x$, then $(zI - T)^{-1} x = (z - \lambda)^{-1} x$ (Exercise 9.3.1) and so

$$P_\lambda x = \frac{1}{2\pi i} \int_\gamma f_\lambda(z)\, (zI - T)^{-1} x\, dz$$

$$= \frac{1}{2\pi i}\left(\int_\gamma f_\lambda(z)\,(z-\lambda)^{-1}\,dz\right)x = f_\lambda(\lambda)\,x = x.$$

This shows that $P_\lambda\mathcal{X} = X_\lambda$. □

The Cayley–Hamilton Theorem is a well-known result in linear algebra. We can use the holomorphic functional calculus to prove it.

**9.3.5. Theorem. (Cayley–Hamilton)**

*Let $n \in \mathbb{N}$ and $T \in M_n(\mathbb{C})$. If $p_T(z) = \det(T - zI)$ is the characteristic polynomial of $T$, then $p_T(T) = 0$.*

*Proof.* Exercise 9.3.4. □

Besides the independence of the curve from Proposition 9.3.1, there is a second uniqueness to be addressed. We have shown that (9.10) does not depend on the curve $\gamma$, as long as $\gamma$ surrounds $\sigma(a)$. A crucial fact, that is more obvious when dealing with continuous and Borel functional calculus (Corollaries 11.2.6 and 12.4.11) than for holomorphic functional calculus, is that the operator defined in (9.10) only depends on the values of $f$ at $\sigma(a)$. Concretely,

**9.3.6. Proposition.**

*Let $a \in \mathcal{A}$ and $f, g \in H(a)$. If $f|_{\sigma(a)} = g|_{\sigma(a)}$, then $f(a) = g(a)$.*

*Proof.* We only need to prove the equality for a closed curve $\gamma$, as the integral is a sum of integrals over curves. And by Proposition 9.3.1 we can choose $\gamma$ to suit our needs. Let $V$ be open, bounded, with $\sigma(a) \subset V$. Fix $\varepsilon > 0$. As $f - g$ is uniformly continuous on the compact set $\overline{V}$ and $f - g = 0$ on $\sigma(a)$, there exists $\delta > 0$ such that $|f(z) - g(z)| < \varepsilon$ if $z \in \sigma(a) + \delta\mathbb{D}$.

Let $A = \big(\sigma(a) + \frac{\delta}{2}\mathbb{D}\big) \cup B$, where $B$ is the union of the bounded connected components of $\mathbb{C} \setminus \big(\sigma(a) + \frac{\delta}{2}\mathbb{D}\big)$ (that is, $A$ is $\sigma(a)$ fattened by $\delta/2$ and with its "holes" filled). By [Con12, Theorem 14.5.5] the boundary $\partial A$ is a continuous curve $\rho$. By Stone–Weierstrass there exists $\gamma$, differentiable, with $\|\gamma - \rho\|_\infty < \delta/2$. Then $\sigma(a) \subset \gamma$ and $|f(z) - g(z)| < \varepsilon$ for all $z \in \gamma$. By Lemma 9.2.11 the map $z \longmapsto (zI_{\mathcal{A}} - a)^{-1}$ is continuous on $\gamma$, and then by compactness there exists $c > 0$ such that $\|(zI_{\mathcal{A}} - a)^{-1}\| \leq c$ for all $z \in \gamma$. Thus for any $\varphi \in \mathcal{A}^*$ we have $|\varphi((zI_{\mathcal{A}} - a)^{-1})| \leq c\,\|\varphi\|$. Using again the continuity of $\varphi$,

$$|\varphi(f(a) - g(a))| = \left|\int_\gamma (f(z) - g(z))\,\varphi((zI_{\mathcal{A}} - a)^{-1})\,dz\right|$$

$$= \left| \int_0^1 (f(\gamma(t)) - g(\gamma(t)))\, \varphi((\gamma(t)I_{\mathcal{A}} - a)^{-1})\, \gamma'(t)\, dt \right|$$
$$\leq c\, \|\varphi\|\, L(\gamma)\, \varepsilon.$$

By Corollary 5.7.7 we deduce that $\|f(a) - g(a)\| \leq c\, L(\gamma)\, \varepsilon$. As $\varepsilon$ was arbitrary, $f(a) = g(a)$. □

Proposition 9.3.6 is a bit weird in the sense that if $\sigma(a)$ is infinite we automatically get that $f = g$ via Corollary 3.6.2. But when $\sigma(a)$ is finite the Riesz Functional Calculus is still defined in terms of an integral along an infinite curve, so it is not like the argument is simplified in any way.

We state here a direct consequence of the estimate in the proof of Proposition 9.3.6:

**_9.3.7. Proposition._**

*Let $a \in \mathcal{A}$ and $f \in H(a)$, bounded. Then there exists $c > 0$ such that $\|f(a)\| \leq c\, \|f\|_\infty$.*

Another possibly unexpected application of the Riesz functional calculus (as in Proposition 9.3.4) is a crude version of the Jordan form of a matrix ("crude" in the sense that it does not tell the geometric multiplicity for a given eigenvalue).

**_9.3.8. Proposition. (Jordan Form, crude version)_**

*Let $T \in M_n(\mathbb{C})$, with eigenvalues $\{\lambda_1, \ldots, \lambda_r\}$ (no repetitions). There exist matrices $T_j \in M_{n(j)}(\mathbb{C})$, $j = 1, \ldots, r$, with $\sigma(T_j) = \{\lambda_j\}$, and an invertible matrix $S$ such that $T = S^{-1}(T_1 \oplus \cdots \oplus T_r)S$ (where the direct sum here is interpreted as block-diagonal). Each $T_j$ is of the form $T_j = \lambda_j I_{n_j} + N_j$, where $N_j$ is nilpotent (concretely, $N_j^{n_j} = 0$).*

*Proof.* Exercise 9.3.9. □

**9.3.1. Exercises.**

(9.3.1) Let $\mathcal{X}$ be a vector space, $T : \mathcal{X} \to \mathcal{X}$ linear and $\lambda \in \mathbb{C}$, $x \in \mathcal{X}$ such that $Tx = \lambda x$. If $(zI - T)$ is invertible, show that $(zI - T)^{-1}x = (z - \lambda)^{-1}x$.

(9.3.2) Let $\mathcal{A}$ be a Banach algebra and $a \in \mathcal{A}$. One can define the exponential $\exp(a)$ by functional calculus,

$$\exp(a) = \frac{1}{2\pi i} \int_\gamma e^z\, (zI - a)^{-1}\, dz$$

for some curve $\gamma$ that contains $\sigma(a)$. One can also define the exponential via the usual series

$$e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!}.$$

Show that the series makes sense in $\mathcal{A}$, and that $\exp(a) = e^a$.

(9.3.3) Let $\mathcal{A}$ be a Banach algebra and $a, b \in \mathcal{A}$.

(i) Show that if $a, b$ commute (that is, $ab = ba$) then $e^{a+b} = e^a e^b$.

(ii) Show an example where $e^{a+b} \neq e^a e^b$.

(9.3.4) Prove Theorem 9.3.5.

(9.3.5) Let $\mathcal{X}$ be a Banach space, $S, T \in \mathcal{B}(\mathcal{X})$, and $f \in H(T)$. Show that if $ST = TS$, then $Sf(T) = f(T)S$.

(9.3.6) Let $T = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$. Find matrices $A$ and $B$ such that $A^4 = T^3$, and $T = e^B$. Are they unique? *(This exercise is not really related to Riesz Functional Calculus)*

(9.3.7) Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ such that $\sigma(T)$ is not connected.

(i) Show that $\sigma(T) = K_1 \cup K_2$, with $K_1, K_2$ compact and disjoint.

(ii) Show that there exist closed subspaces $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ such that $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ and $S\mathcal{X}_1 \subset \mathcal{X}_1$ and $S\mathcal{X}_2 \subset \mathcal{X}_2$ for all $S \in \mathcal{B}(\mathcal{X})$ such that $ST = TS$.

(iii) Denote $T_1 = T|_{\mathcal{X}_1} \in \mathcal{B}(\mathcal{X}_1)$ and $T_2 = T|_{\mathcal{X}_2} \in \mathcal{B}(\mathcal{X}_2)$. Show that $\sigma(T_1) = K_1$ and $\sigma(T_2) = K_2$.

(iv) Prove that there exists $S : \mathcal{X} \to \mathcal{X}_1 \oplus \mathcal{X}_2$, invertible, with $T = S^{-1}(T_1 \oplus T_2)S$.

(9.3.8) Let $\mathcal{A}$ be the Banach algebra $C(X)$ for compact Hausdorff $X$. Let $f \in \mathcal{A}$ and $g \in H(f)$. Show that $g(f) = g \circ f$.

(9.3.9) Prove Proposition 9.3.8.

## 9.4. Adjoints

The notion of the adjoint of an operator is natural and it will show itself very useful.

**9.4.1. Definition.**

*Given $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$, the* ***adjoint*** *of $T$ is the operator $T^* : \mathcal{Y}^* \to \mathcal{X}^*$, given by*

$$(T^*g)(x) = g(Tx), \quad g \in \mathcal{Y}^*,\ x \in \mathcal{X}.$$

The first-time reader should pay attention to the fact that in Definition 9.4.1 the $*$ represents two different things, namely the dual of a normed space, and the adjoint of an operator. There should be no confusion, though, as duals are spaces and adjoints are operators.

Basic properties are that $T^*$ is bounded, that $\|T^*\| = \|T\|$ (Exercise 9.4.3), and that the map $T \longmapsto T^*$ is linear and anti-multiplicative with respect to composition (Exercise 9.4.2). There is a subtle difference between this notion of adjoint and the adjoint for an operator on a Hilbert space, as we will discuss in Section 10.1. When $\mathcal{X} = \mathbb{C}^n$, if we see $T$ as an $n \times n$ complex matrix and we assume $\mathcal{X}^* = \mathcal{X}$ via the obvious dual basis, the map $T^*$ can be seen as the transpose of $T$.

In Proposition 9.1.2 we showed that weak-to-weak continuous operators are bounded. Now we can see that adjoints are precisely the weak$^*$-continuous maps:

***9.4.2. Proposition.***

*Let $\mathcal{X},\mathcal{Y}$ be normed spaces. A linear map $S : \mathcal{Y}^* \to \mathcal{X}^*$ is weak$^*$ continuous if and only if $S = T^*$ for some $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$*

*Proof.* If $S = T^*$, and $f_j \to f$ weak$^*$ in $\mathcal{Y}^*$, then for any $x \in \mathcal{X}$

$$Sf_j(x) = f_j(Tx) \to f(Tx) = T^*f(x) = Sf(x).$$

So $S$ is weak$^*$-continuous.

Conversely, assume that $S$ is weak$^*$-continuous. We know what our $T$ should satisfy if it exists: $Sf(x) = f(Tx)$. So let us use this to define $T$.

For any $x \in \mathcal{X}$, consider the functional on $\mathcal{Y}^*$ given by $\alpha_x : f \longmapsto Sf(x)$. By hypothesis this is weak$^*$-continuous; the weak$^*$-continuous functionals are bounded (Exercise 7.2.12), so $\alpha_x \in Y^{**}$. But, also, a weak$^*$-continuous functional on the dual is in the predual by Proposition 7.2.10. So there exists $y \in \mathcal{Y}$ with $\alpha_x(f) = f(y)$ for all $f \in \mathcal{Y}^*$. Define $Tx = y$. Note that $y$ is unique because $\mathcal{Y}^*$ separates points in $\mathcal{Y}$. From there we deduce that $T$ is linear. Finally, we have that $S$ is bounded; indeed, if $f_j \to f$ in $\mathcal{Y}^*$ and $Sf_j \to g$ in $X^*$ (both in norm), then $f_j \to f$ in the weak$^*$ topology, so $Sf_j \to Sf$ weak$^*$; thus $Sf = g$ and $S$ is bounded by the Closed Graph Theorem 6.3.12 (note that both $\mathcal{Y}^*$ and $\mathcal{X}^*$ are Banach). Then

$$\|Tx\| = \sup\{|f(Tx)| :\ f \in \mathcal{Y}^*,\ \|f\| = 1\}$$

$$= \sup\{|Sf(x)| : \ f \in \mathcal{Y}^*, \ \|f\| = 1\} \leq \|S\| \, \|x\|.$$

So $T$ is bounded with $\|T\| \leq \|S\|$. □

**9.4.3. Definition.**

*An operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is* ***bounded below*** *if there exists $c > 0$ such that*

$$\|x\| \leq c\|Tx\|, \qquad x \in \mathcal{X}.$$

It is obvious that if $T$ is bounded below, it is injective. And a bounded below operator has closed range (Exercise 9.4.1). The converse is not true:

*9.4.4. Example.* Let $\mathcal{X} = \mathcal{Y} = \ell^\infty(\mathbb{N})$. Define

$$Sx = (x_1, 0, 0, \ldots).$$

Then $S$ has closed range (it is a rank-one operator) and $Te_2 = 0$, so $T$ is not bounded below. Here is a surjective example:

$$Tx = (x_1 + x_2, x_3 + x_4/2, x_5 + x_6/3, \ldots).$$

Then $T$ is surjective (so it has closed range); indeed, given any $y \in \mathcal{X}$,

$$y = T(y_1, 0, y_2, 0, y_3, \ldots).$$

But it is not bounded below: given the element $e_{2n}$ in the canonical basis, we have $\|e_{2n}\| = 1$ and

$$\|Te_{2n}\| = \tfrac{1}{n} = \tfrac{1}{n}\,\|e_{2n}\|.$$

The following technical lemma will be used several times. As mentioned above, a bounded below operator has closed range. The lemma shows that the converse "almost" holds, and the actual converse holds when $T$ is injective.

***9.4.5. Lemma.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range. Then there exists $c > 0$ such that for every $y \in T\mathcal{X}$ there exists $x \in \mathcal{X}$ with $Tx = y$ and $\|x\| \leq c\,\|Tx\|$. If $T$ is also injective, then it is bounded below.*

*Proof.* By considering the surjective operator $T : \mathcal{X} \to \operatorname{ran} T$, the Open Mapping Theorem (Theorem 6.3.5) gives us that there exists $r > 0$ such that $B_r^{\mathcal{Y}}(0) \subset TB_1^{\mathcal{X}}(0)$. Given any $y \in T\mathcal{X}$, we have $\frac{r}{2\|y\|}\, y \in B_r^{\mathcal{Y}}(0)$. So there exists $x_0 \in B_1^{\mathcal{X}}(0)$ with $\frac{r}{2\|y\|}\, y = Tx_0$. Then

$$\|x_0\| < 1 = \left\| \frac{y}{\|y\|} \right\| = \frac{2}{r}\, \|Tx_0\|.$$

We take $x = \frac{2\|y\|}{r}\, x_0$, $c = 2/r$ and we are done.

When $T$ is also injective, given $x \in \mathcal{X}$ by the above there exists $x' \in \mathcal{X}$ with $\|x'\| \le c\|Tx'\|$ and $Tx' = Tx$. But as $T$ is injective, $x' = x$, so $\|x\| \le c\|Tx\|$ and $T$ is bounded below. □

The next result lists several relations between $T$ and $T^*$. We will use polars and prepolars as considered in Section 7.3.

***9.4.6. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then*

*(i) if $T^*$ is invertible, then $\operatorname{ran} T$ is closed;*

*(ii)*
$$\ker T^* = (\operatorname{ran} T)^o. \tag{9.11}$$

*(iii) if $\operatorname{ran} T$ is closed, then*
$$\operatorname{ran} T^* = (\ker T)^o; \tag{9.12}$$

*(iv)*
$$\overline{\operatorname{ran}} T = (\ker T^*)_o, \tag{9.13}$$
$$(\operatorname{ran} T^*)_o = \ker T. \tag{9.14}$$

*(v) $T$ is injective if and only if $\operatorname{ran} T^*$ is weak*-dense;*

*(vi) $T^*$ is injective if and only if $\operatorname{ran} T$ is dense;*

*(vii) $\operatorname{ran} T$ is closed if and only if $\operatorname{ran} T^*$ is closed;*

*(viii) $T$ is bounded below if and only if $T^*$ is surjective.*

*Proof.* (i) Assume that $T^*$ is invertible; so there exists $W \in \mathcal{B}(\mathcal{X}^*, y^*)$ with $WT^* = I_{\mathcal{Y}^*}$, $T^*W = I_{\mathcal{X}^*}$. Suppose that $\{Tx_n\} \subset \mathcal{Y}$ is Cauchy. Then by Corollary 5.7.7

$$\begin{aligned}
\|x_n - x_m\| &= \sup\{|f(x_n - x_m)| : \ f \in \mathcal{X}^*, \ \|f\| = 1\} \\
&= \sup\{|WT^*f(x_n - x_m)| : \ \|f\| = 1\} \\
&\le \|W\| \, \sup\{|f(Tx_n - Tx_m)| : \ \|f\| = 1\} \\
&\le \|W\| \, \|Tx_n - Tx_m\|.
\end{aligned}$$

So $\{x_n\}$ is Cauchy. If $x = \lim x_n$, then $Tx = \lim Tx_n$, and $\operatorname{ran} T$ is closed.

(ii) We have, for $g \in \mathcal{Y}^*$, $x \in \mathcal{X}$,

$$(T^*g)x = g(Tx).$$

If $g \in \ker T^*$, the above equality shows that $g \in (\operatorname{ran} T)^o$. And vice versa.

(iii) Suppose first that $f \in \operatorname{ran} T^*$. Then $f = T^*g$ for some $g \in \mathcal{Y}^*$. If $x \in \ker T$,

$$f(x) = (T^*g)x = g(Tx) = g(0) = 0.$$

Conversely, suppose that $f(x) = 0$ for all $x \in \ker T$. On $T\mathcal{X}$, define $\phi : T\mathcal{X} \to \mathbb{C}$ as $\phi(Tx) = f(x)$. This is well-defined, because $f = 0$ on $\ker T$. Fix $x \in \mathcal{X}$. Since $\operatorname{ran} T$ is closed, by Lemma 9.4.5 there exists $c > 0$ and $x' \in \mathcal{X}$ with $Tx' = Tx$ and $\|x'\| \leq c\|Tx\|$. Then

$$|\phi(Tx)| = |\phi(Tx')| = |f(x')| \leq \|f\|\,\|x'\| \leq c\|f\|\,\|Tx\|.$$

So $\phi$ is bounded. By Hahn–Banach (Corollary 5.7.6), there exists $g \in \mathcal{Y}^*$ such that $g|_{T\mathcal{X}} = \phi$, and $\|g\| = \|\phi\|$. Now $(T^*g)x = g(Tx) = \phi(Tx) = f(x)$, and so $f \in \operatorname{ran} T^*$.

(iv) from (9.11) we get $(\ker T^*)_o = [(\operatorname{ran} T)^o]_o$. As $\operatorname{ran} T$ is convex and balanced, being a subspace, Proposition 7.3.4 gives us (9.13).

If $x \in (\operatorname{ran} T^*)_o$, then $0 = (T^*g)x = g(Tx)$ for all $g \in \mathcal{Y}^*$. By Hahn–Banach (Corollary 5.7.7), $Tx = 0$ and so $x \in \ker T$. Conversely, if $Tx = 0$ the same computations show that $(T^*g)x = 0$ for all $g \in \mathcal{Y}^*$, so $x \in (\operatorname{ran} T^*)_o$ and (9.14) is established.

(v) This follows directly from (9.14) and (ii) in Corollary 7.3.5.

(vi) This follows directly from (9.13) and (i) in Corollary 7.3.5.

(vii) if $\operatorname{ran} T$ is closed it follows directly from (9.12) that $\operatorname{ran} T^*$ is closed, since polars are closed. Conversely, if $\operatorname{ran} T^*$ is closed, let $S : \mathcal{Y}^* \to \operatorname{ran} T^*$ be given by $Sg = T^*g$. By hypothesis, $S$ is a surjective map between Banach spaces. By (vi), $S^*$ is injective. Since $\operatorname{ran} S$ is closed, so is $\operatorname{ran} S^*$ by the first part of the proof. This means that $S^*$ is injective with closed range; by Lemma 9.4.5, it is bounded below: there exists $c > 0$ such that $\|\phi\| \leq c\|S^*\phi\|$ for all $\phi \in (\operatorname{ran} T^*)^*$. Given $x \in \mathcal{X}$, we have $\hat{x} \in \mathcal{X}^{**}$; as $\operatorname{ran} T^* \subset \mathcal{X}^*$, we can consider $\hat{x} \in (\operatorname{ran} T^*)^*$ by restriction. Then, for any $g \in \mathcal{Y}^*$,

$$(S^*\hat{x})g = \hat{x}(Sg) = \hat{x}(T^*g) = (T^*g)x = g(Tx).$$

Via Corollary 5.7.7,

$$\|S^*\hat{x}\| = \sup\{|(S^*\hat{x})g| : \ \|g\| = 1\} = \sup\{|g(Tx)| : \ \|g\| = 1\} = \|Tx\|.$$

Using that $S^*$ is bounded below,

$$\|x\| = \|\hat{x}\| \leq c\|S^*\hat{x}\| = c\|Tx\|,$$

so $T$ is bounded below, and thus it has closed range.

(viii) If $T$ is bounded below, then $\operatorname{ran} T$ is closed. Consider $T_0 : \mathcal{X} \to \operatorname{ran} T$ with $T_0x = Tx$. Since $T_0$ is bounded below and surjective, it is invertible. That is, there exists $W : \operatorname{ran} T \to \mathcal{X}$ with $WT_0 = I_{\mathcal{X}}$, $T_0W = I_{\operatorname{ran} T}$. In particular, $T_0^*W^* = (WT_0)^* = I_{\mathcal{X}}^* = I_{\mathcal{X}^*}$ (where we use Exercise 9.4.2), so $T_0^*$ is surjective. Given $f \in \mathcal{X}^*$, there exists $g_0 \in (\operatorname{ran} T)^*$ with $T_0^* g_0 = f$. We use Hahn–Banach (Corollary 5.7.6) to extend $g_0$ to $g \in \mathcal{Y}^*$. Then, for any $x \in \mathcal{X}$,

$$(T^* g)x = g(Tx) = g_0(T_0 x) = (T_0^* g_0)x = f(x).$$

Hence $T^* g = f$, and $T^*$ is surjective.

Conversely, if $T^*$ is surjective, then $T$ is injective by (v), and $\operatorname{ran} T$ is closed by (vii). Lemma 9.4.5 then shows that $T$ is bounded below. □

*9.4.7. Remark*. We cannot expect the converse of (i) in Proposition 9.4.6 to hold. The unilateral shift, discussed in Example 9.5.10 and beyond, provides a counterexample.

*9.4.8. Remark*. If $\mathcal{X}$ is reflexive, then $T^{**} = T$ (Exercise 9.4.6) and the roles of $T$ and $T^*$ become symmetric in Proposition 9.4.6.

When $\operatorname{ran} T$ is closed, Proposition 9.4.6 takes a much simpler form:

***9.4.9. Corollary.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, with $\operatorname{ran} T$ closed. Then $\operatorname{ran} T^*$ is closed, and*

*(i) $T$ is injective if and only if $T^*$ is surjective;*

*(ii) $T$ is surjective if and only if $T^*$ is injective.*

*Proof.* Since $\operatorname{ran} T$ is closed, $\operatorname{ran} T^*$ is closed by (vii) in Proposition 9.4.6.

If $T$ is injective, as it has closed range it is bounded below by Lemma 9.4.5. Then $T^*$ is surjective by (viii) in Proposition 9.4.6. If $T^*$ is surjective, then $T$ is injective by (v) in Proposition 9.4.6. Finally, since $\operatorname{ran} T$ is closed, $T$ is surjective if and only if $T^*$ is injective by (vi) in Proposition 9.4.6. □

The following result will prove useful.

***9.4.10. Proposition.***

*Let $\mathcal{X}$ be a Banach space and $\mathcal{Y}$ a normed space. Then the unit ball of $\mathcal{B}(\mathcal{X}, \mathcal{Y}^*)$ is compact for the topology of pointwise weak*-convergence.*

*Proof.* We will show that $\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$ is naturally a dual, and that the pointwise weak* topology is the weak* topology of such identification. Let

$$\mathcal{D}=\overline{\operatorname{span}}^{\|\cdot\|}\{g_{x,y}:\mathcal{B}(\mathcal{X},\mathcal{Y}^*)\to\mathbb{C}:\ x\in\mathcal{X},\ y\in\mathcal{Y},\ g_{x,y}(T)=(Tx)y\}.$$

It follows that $\|g_{x,y}\|=\|x\|\,\|y\|$ (Exercise 9.4.8). We will show that $\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$ is isometrically isomorphic to $\mathcal{D}^*$, via the duality $\langle\hat{T},g\rangle=g(T)$. The map $\pi:\mathcal{B}(\mathcal{X},\mathcal{Y}^*)\to\mathcal{D}^*$ given by $\pi:T\longmapsto\hat{T}$ is obviously linear. For any $T\in\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$ and $g\in\mathcal{D}$,

$$|\hat{T}(g)|=|g(T)|\leq\|g\|\,\|T\|,$$

and so $\|\hat{T}\|\leq\|T\|$. Conversely, fix $\varepsilon>0$. Then there exists $x\in\mathcal{X}$ with $\|x\|=1$ and $\|T\|\leq\|Tx\|+\varepsilon/2$. There also exists $y\in\mathcal{Y}$ with $\|y\|=1$ and $\|Tx\|\leq|(Tx)y|+\varepsilon/2$. Then

$$\|T\|\leq|(Tx)y|+\varepsilon=|g_{x,y}(T)|+\varepsilon=|\hat{T}(g_{x,y})|+\varepsilon\leq\|\hat{T}\|\,\|g_{x,y}\|+\varepsilon=\|\hat{T}\|+\varepsilon.$$

As $\varepsilon$ was arbitrary, we conclude that $\|\hat{T}\|=\|T\|$. It remains to show that $\pi$ is surjective. Let $f\in\mathcal{D}^*$, and fix $x\in\mathcal{X}$ with $\|x\|=1$. The map $y\longmapsto f(g_{x,y})$ is a bounded linear functional on $\mathcal{Y}$; so there exists a unique $T_x\in\mathcal{Y}^*$ with $f(g_{x,y})=T_x(y)$. The uniqueness of $T_x$, together with the bilinearity of $g_{x,y}$, guarantee that $T:x\longmapsto T_x$ is linear. Also,

$$\begin{aligned}\|Tx\|=\|T_x\|&=\sup\{|T_x(y)|:\ y\in\mathcal{Y}^*,\ \|y\|=1\}\\&=\sup\{|f(g_{x,y})|:\ y\in\mathcal{Y}^*,\ \|y\|=1\}\\&\leq\sup\{\|f\|\,\|g_{x,y}\|:\ \|y\|=1\}\leq\|f\|\,\|x\|.\end{aligned}$$

So $\|T\|\leq\|f\|$, and $T\in\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$. Now

$$\hat{T}(g_{x,y})=g_{x,y}(T)=(Tx)y=f(g_{x,y}).$$

As the span of the $g_{x,y}$ is dense in $\mathcal{D}$, we have shown that $\hat{T}=f$. So $\pi$ is surjective and we are done showing that $\mathcal{D}^*\simeq\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$.

Let $\{T_j\}$ be a net in the unit ball of $\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$ such that $T_j\to T$ pointwise weak*. That is, $T_jx\to Tx$ weak* for all $x\in\mathcal{X}$. Then

$$\hat{T}_j(g_{x,y})=(T_jx)y\to(Tx)y=\hat{T}(g_{x,y}).$$

As the $g_{x,y}$ span a dense subset of $\mathcal{D}$, we have shown that $\hat{T}_j\to\hat{T}$ weak* in $\mathcal{D}^*$. Conversely if $\hat{T}_j\to\hat{T}$ weak*, undoing the above steps we get that $T_j\to T$ pointwise weak*.

Now Banach–Alaoglu (Theorem 7.2.13) tells us that the closed unit ball of $\mathcal{B}(\mathcal{X},\mathcal{Y}^*)$ is compact in the weak*-topology as a the dual of $\mathcal{D}$, which we have shown is the pointwise weak* topology. $\square$

With knowledge about adjoints, we can prove the following result related to Theorem 7.4.8. We will need to work with real linear maps. It would be an interesting rabbit hole to go down and check that all results quoted work in the real case; but the main point is that Hahn–Banach works in the real case (in fact, we first proved a real version before moving onto a complex one) and, beyond that, one does not really use that scalars are complex. Note that linearity of $T$ is not required in the statement of the theorem.

**9.4.11. Theorem. (Banach–Stone)**

*Let $X, Y$ be locally compact Hausdorff spaces and $T : C_0(X) \to C_0(Y)$ a surjective (not necessarily linear) isometry. Then there exists a homeomorphism $\tau : Y \to X$ and $g, h \in C(Y)$ such that $|h| = 1$ and*

$$(Tf)(y) = g(y) + h(y)\, f(\tau(y)), \qquad f \in C(X),\ y \in Y.$$

*Proof.* By Proposition 5.1.18 and Lemma 5.1.17 we may assume that $T - T0$ is real linear. In that case the associated real linear transform will also be a surjective isometry. So we may assume without loss of generality that $T$ is real linear, and then use $g = T0$ in the final statement. Because $T$ is isometric and surjective,

$$\begin{aligned}\|T^*\varphi\| &= \sup\{|(T^*\varphi)f| : \ f \in C(X),\ \|f\| = 1\}\\ &= \sup\{|(\varphi(Tf)| : \ f \in C(X),\ \|f\| = 1\}\\ &= \sup\{|(\varphi(g)| : \ g \in C(Y),\ \|g\| = 1\} = \|\varphi\|.\end{aligned}$$

So $T^*$ is isometric, too, and by Corollary 9.4.9 it is surjective. So both $T$ and $T^*$ are invertible (we don't need the inverse mapping theorem for this, as we already know they are isometries). With $T^*$ being an isometric bijection, it maps the unit ball of $C_0(Y)^*$ onto the unit ball of $C_0(X)^*$. From this and the real linearity of $T^*$ we can obtain that $T^*$ maps $\operatorname{Ext}\overline{B_1^{C_0(Y)^*}(0)}$ onto $\operatorname{Ext}\overline{B_1^{C_0(X)^*}(0)}$ (Exercise 7.5.19). We know by Exercise 7.5.16 that these are the Dirac measures, multiplied by complex numbers of absolute value 1. So, for each $y \in Y$ there exists a unique $\tau(y) \in X$ and a unique number $h(y) \in \mathbb{T}$ such that

$$T^*\delta_y = h(y)\,\delta_{\tau(y)}.$$

By Proposition 9.4.2, the map $T^*$ is weak*-continuous. If $\{y_j\} \subset Y$ is a net with $y_j \to y$, then

$$\begin{aligned}\lim_j h(y_j) &= \lim_j h(y_j)\delta_{y_j}(1) = \lim_j (T^*\delta_{y_j})(1)\\ &= (T^*\delta_y)(1) = h(y)\,\delta_{\tau(y)}(1) = h(y).\end{aligned}$$

So $h$ is continuous. Since $h$ is continuous and never zero, we can now write for $f \in C_0(X)$

$$\lim_j \delta_{\tau(y_j)} f = \lim_j h(y_j)^{-1}\, h(y_j)\delta_{\tau(y_j)} f = \lim_j h(y_j)^{-1}\,(T^*\delta_{y_j}) f$$

$$= h(y)^{-1}\,(T^*\delta_y)f = h(y)^{-1}\,h(y)\delta_{\tau(y)}f = \delta_{\tau(y)}f.$$

That is, $\lim_j f(\tau(y_j)) = f(\tau(y))$ for all $f \in C_0(X)$. By Corollary 2.6.7, $\tau(y_j) \to \tau(y)$, and so $\tau$ is continuous. If $y_1, y_2 \in Y$ are distinct, then $\delta_{y_1} \neq \delta_{y_2}$; this implies that $h(y_1)^{-1}\delta_{y_1} \neq h(y_2)^{-1}\delta_{y_2}$. For we can always use Urysohn's Lemma to find a continuous function $g$ with $g(y_1) = 1$, $g(y_2) = 0$, leading to a contradiction. By the injectivity of $T^*$, this means

$$\delta_{\tau(y_1)} = T^*(h(y_1)^{-1}\delta_{y_1}) \neq T^*(h(y_2)^{-1}\delta_{y_2}) = \delta_{\tau(y_2)}$$

and then $\tau(y_1) \neq \tau(y_2)$, so $\tau$ is injective. Now fix $x \in X$. Since $T^*$ is surjective from $\operatorname{Ext}\overline{B_1^{C(Y)^*}(0)}$ onto $\operatorname{Ext}\overline{B_1^{C(X)^*}(0)}$, there exists $y \in Y$ and $\alpha \in \mathbb{T}$ such that $T^*(\alpha\,\delta_y) = \delta_x$. Then

$$\delta_x = T^*(\alpha\,\delta_y) = \alpha\,h(y)\,\delta_{\tau(y)}.$$

This forces $\alpha = h(y)^{-1}$ and $\tau(y) = x$. So $\tau$ is surjective. Now we know that $\tau : Y \to X$ is a continuous bijection. It remains to show that $\tau^{-1}$ is continuous. If $\tau(y_j) \to \tau(y)$, then $\delta_{\tau(y_j)} \to \delta_{\tau(y)}$ weak$^*$ since they evaluate on continuous functions. Then

$$T^*(h(y_j)^{-1}\delta_{y_j}) = \delta_{\tau(y_j)} \to \delta_{\tau(y)} = T^*(h(y)^{-1}\delta_y).$$

As $T^*$ is bijective and bicontinuous, we get that $h(y_j)^{-1}\delta_{y_j} \to h(y)^{-1}\delta_y$ weak$^*$. That is,

$$h(y_j)^{-1}f(y_j) \to h(y)^{-1}f(y), \qquad f \in C_0(Y).$$

Since $|h| = 1$, when $f$ runs over all of $C_0(Y)$ we have that $h^{-1}f$ also runs over all of $C_0(Y)$. Then $y_j \to y$ by Corollary 2.6.7 and $\tau$ is a homeomorphism.

Finally, given $f \in C(X)$ and $y \in Y$,

$$(Tf)(y) = \delta_y(Tf) = (T^*\delta_y)f = h(y)\,\delta_{\tau(y)}f = h(y)\,f(\tau(y)). \qquad \square$$

### 9.4.1. Exercises.

(9.4.1) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that if $T$ is bounded below, then $\operatorname{ran} T$ is closed.

(9.4.2) Let $T \in \mathcal{B}(\mathcal{X})$. Show that the map $T \longmapsto T^*$ is linear and anti-multiplicative.

(9.4.3) Given $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, show that its adjoint $T^*$ is bounded, and that $\|T^*\| = \|T\|$.

(9.4.4) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ be an isometric isomorphism. Show that $T^*$ is an isometric isomorphism.

(9.4.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that

$$\|T^{**}\hat{x}\| = \|Tx\|, \qquad x \in \mathcal{X}.$$

(9.4.6) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that if $\mathcal{X}$ is reflexive, then $T^{**} = T$.

(9.4.7) Use Proposition 9.4.2 to give an alternative proof of Proposition 7.2.10.

(9.4.8) In the context of the proof of Proposition 9.4.10, show that
$$\|g_{x,y}\| = \|x\|\,\|y\|.$$
(9.4.9) Prove (v) and (vi) in Proposition 9.4.6 without using polars nor prepolars.

(9.4.10) The following is a "counterexample" to Proposition 9.4.2. Find the mistake.

> *Take $\mathcal{Y} = c_{00} \subset \mathcal{X} = \ell^1(\mathbb{N})$; so we consider the 1-norm on $\mathcal{Y}$. Because $\mathcal{Y}$ is dense in $\ell^1(\mathbb{N})$, we have $\mathcal{X}^* = \mathcal{Y}^* = \ell^\infty(\mathbb{N})$.*
>
> *Define $S : \mathcal{Y}^* \to \mathcal{X}^*$, that is $S : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$, by*
> $$Sw = \Big(\sum_n \frac{w(n)}{n^2}, 0, 0, \ldots\Big).$$
> *If $w_j \to 0$ weak*, this means that $\sum_n w_j(n)x(n) \to 0$ for all $x \in \mathcal{X}$. In particular $\sum_n \frac{w_j(n)}{n^2} \to 0$, and it follows $S$ is weak*-weak* continuous. If we had $S = T^*$, with $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ this would mean that, for each $w \in \ell^\infty(\mathbb{N})$ and $x \in \mathcal{X}$,*
> $$(Sw)x = w(Tx).$$
> *This translates to*
> $$\sum_n \frac{w(n)x(1)}{n^2} = \sum_n w(n)\,(Tx)(n).$$
> *As this should work for all $w \in \ell^\infty$, it follows that we need*
> $$Tx = \left(\frac{x(1)}{n^2}\right)_n.$$
> *But then $Tx \notin \mathcal{Y}$ for any nonzero $x$, and so $T \notin \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

(9.4.11) Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be Banach spaces, $S : \mathcal{Y} \to \mathcal{Z}$ linear, and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range. Use Lemma 9.4.5 to show that if $ST \in \mathcal{B}(\mathcal{X}, \mathcal{Z})$ then $S$ is bounded.

(9.4.12) We use notation from Section 9.3. Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$, and $f \in H(T)$. Show that $f(T^*) = f(T)^*$.

## 9.5. The Spectrum of a Linear Operator

The results in Section 9.2 apply in particular to the Banach algebra $\mathcal{B}(\mathcal{X})$ of bounded linear operators on a Banach space $\mathcal{X}$. But the fact that the elements in our Banach algebra are operators allow us to consider the spectrum in more detail. We only consider the case where $\mathcal{X}$ is Banach, for otherwise $\mathcal{B}(\mathcal{X})$ is not a Banach algebra, and since the spectrum deals with invertibility, problems arise; we discussed this briefly in Examples 9.2.7 with the case of $\mathbb{C}[x]$.

Our first goal is to relate the spectrum of $T$ and $T^*$.

***9.5.1. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X})$. Then $\sigma(T^*) = \sigma(T)$.*

*Proof.* Since $(T-\lambda I_{\mathcal{X}})^* = T^* - \lambda I_{\mathcal{X}^*}$, it is enough to show that $T$ is invertible if and only if $T^*$ is invertible (make sure you understand why!). If $T$ is invertible, there exists $S$ with $TS = ST = I_{\mathcal{X}}$. Then $T^*S^* = (ST)^* = I_{\mathcal{X}}^* = I_{\mathcal{X}^*}$, and similarly $S^*T^* = I_{\mathcal{X}^*}$; so $T^*$ is invertible.

Conversely, assume that $T^*$ is invertible; so there exists $W \in \mathcal{B}(\mathcal{X}^*)$ with $WT^* = T^*W = I_{\mathcal{X}^*}$. By Proposition 9.4.6, we have that $\operatorname{ran} T$ is dense and closed, so $T$ is surjective. As $T^*$ is surjective, $T$ is injective, again by Proposition 9.4.6. So $T$ is bijective and bounded. By the Inverse Mapping Theorem 6.3.6, $T$ is invertible. □

The above result is one that holds for operators on Banach spaces and not for operators on Hilbert spaces, due to the fact that the definition of the adjoint is not exactly the same in the latter case.

When $T \in \mathcal{B}(\mathcal{X})$ and $\mathcal{X}$ is finite dimensional, $T$ is injective if and only if it is surjective; and its range is always closed. This amounts to a single point of failure for invertibility: $T$ is not invertible if and only if $T$ is not surjective if and only if $T$ is not injective. Thus the spectrum of $T$ consists precisely of the eigenvalues: $T - \lambda I$ is not invertible if and only if $\ker(T - \lambda I) \neq \{0\}$, and its elements are eigenvectors for $\lambda$.

When $\mathcal{X}$ is infinite dimensional, the three "points of failure" to be invertible can and do occur separately. Because of that, we recognize different classes of elements of the spectrum. Thus, we define (the list is not exhaustive, although it contains the most common classes)

**9.5.2. Definition.**

- *The* ***point spectrum*** *of $T$ (eigenvalues):*
$$\sigma_p(T) = \{\lambda : \ \ker(T - \lambda I) \neq \{0\}\};$$
- *the* ***approximate point spectrum*** *of $T$:*
$$\sigma_{ap}(T) = \{\lambda : \ \exists\{x_n\} \subset \mathcal{X}, \ \|x_n\| = 1, \ Tx_n - \lambda x_n \to 0\};$$
- *the* ***residual spectrum*** *of $T$:*
$$\sigma_r(T) = \{\lambda : \ \lambda \notin \sigma_p(T), \ \ \textit{and } T - \lambda I \textit{ does not have dense range}\};$$
- *the* ***continuous spectrum*** *of $T$:*
$$\sigma_c(T) = \{\lambda \colon \lambda \notin \sigma_p(T), \textit{ and } T - \lambda I \textit{ has dense range but not surjective}\}.$$

These parts of the spectrum are not all disjoint. We have, for instance,
$$\sigma_p(T) \subset \sigma_{ap}(T), \tag{9.15}$$
and
$$\sigma_c(T) = \sigma_{ap}(T) \setminus (\sigma_r(T) \cup \sigma_p(T)). \tag{9.16}$$
Also, $\sigma_p(T)$, $\sigma_r(T)$, and $\sigma_c(T)$ are mutually disjoint and
$$\sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T), \tag{9.17}$$
$$\sigma(T) = \sigma_{ap}(T) \cup \sigma_r(T) \tag{9.18}$$
(Exercises 9.5.9 and 9.5.11). But it is not necessarily true that $\sigma_{ap}(T) \cap \sigma_r(T) = \varnothing$ (Exercise 9.5.12).

We will frequently use the fact that (9.17) is a disjoint union.

***9.5.3. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X})$. Then*
$$\sigma_p(T) \subset \sigma_p(T^{**}), \tag{9.19}$$
$$\sigma_r(T) \subset \sigma_p(T^*) \subset \sigma_r(T) \cup \sigma_p(T), \tag{9.20}$$
*When $\mathcal{X}$ is reflexive, we have*
$$\sigma_p(T) = \sigma_p(T^{**}), \tag{9.21}$$
$$\sigma_r(T^*) \subset \sigma_p(T) \subset \sigma_r(T^*) \cup \sigma_p(T^*), \tag{9.22}$$
*and*
$$\sigma_c(T) = \sigma_c(T^*). \tag{9.23}$$

*Proof.* Let $\lambda \in \sigma_p(T)$; so there exists nonzero $x \in X$ with $Tx = \lambda x$. For any $g \in \mathcal{X}^*$,
$$(T^{**}\hat{x})g = \hat{x}(T^*g) = (T^*g)(x) = g(Tx) = g(\lambda x) = (\lambda\hat{x})g.$$

So $T^{**}\hat{x} = \lambda\hat{x}$, showing that $\lambda \in \sigma_p(T^{**})$.

If $\lambda \in \sigma_r(T)$, then $T - \lambda I_{\mathcal{X}}$ does not have dense range. By Proposition 9.4.6, $T^* - \lambda I_{\mathcal{X}^*}$ is not injective; that is, $\lambda \in \sigma_p(T^*)$. If $\lambda \in \sigma_p(T^*)$, then $T^* - \lambda I_{\mathcal{X}^*}$ is not injective; again by Proposition 9.4.6, $\operatorname{ran}(T - \lambda I)$ is not dense. So if $\lambda \notin \sigma_p(T)$, we get that $\lambda \in \sigma_r(T)$ and (9.20) is established.

When $\mathcal{X}$ is reflexive, $T^{**} = T$ (Exercise 9.4.6), and we may apply the arguments above to $T^*$. If $\lambda \in \sigma_c(T)$, by Proposition 9.5.1 we have that $\lambda \in \sigma(T^*)$. Since $\lambda \notin \sigma_r(T) \cup \sigma_p(T)$, it follows from (9.20) that $\lambda \notin \sigma_p(T^*)$. Thus $\lambda \in \sigma_c(T^*) \cup \sigma_r(T^*)$. If we had $\lambda \in \sigma_r(T^*)$, by (9.20) we would have $\lambda \in \sigma_p(T^{**}) = \sigma_p(T)$, contradicting that $\lambda \in \sigma_c(T)$. Hence $\sigma_c(T) \subset \sigma_c(T^*)$, and now the reflexivity gives us the reverse inclusion and $\sigma_c(T) = \sigma_c(T^*)$. □

The next two results will help us with some spectral calculations.

***9.5.4. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X})$, $\lambda \in \mathbb{C}$. The following statements are equivalent:*

*(i)* $\lambda \notin \sigma_{ap}(T)$;

*(ii)* $T - \lambda I$ *is bounded below;*

*(iii)* $\ker(T - \lambda I) = \{0\}$ *and* $\operatorname{ran}(T - \lambda I)$ *is closed;*

*Proof.* Since $\lambda \in \sigma_{ap}(T)$ if and only if $0 \in \sigma_{ap}(T - \lambda I)$, we may assume without loss of generality that $\lambda = 0$.

(i) $\Longrightarrow$ (ii) Suppose that $T$ is not bounded below. Then for each $n \in \mathbb{N}$ there exists $x_n' \in \mathcal{X}$ with $\|Tx_n'\| < \frac{1}{n}\|x_n'\|$. Let $x_n = \frac{x_n'}{\|x_n'\|}$. Then $\|x_n\| = 1$ for all $n$, and $\|Tx_n\| < \frac{1}{n} \to 0$. So $\lambda \in \sigma_{ap}(T)$.

(ii) $\Longrightarrow$ (iii) The injectivity of $T$ is automatic from being bounded below, so $\ker T = \{0\}$. And $\operatorname{ran} T$ is closed by Exercise 9.4.1.

(iii) $\Longrightarrow$ (i) Since $T$ is injective and $\operatorname{ran} T$ is closed, the map $T : \mathcal{X} \to \operatorname{ran} T$ is bijective and bounded linear map between Banach spaces. By the Inverse Mapping Theorem 6.3.6 we get that $T$ is invertible. So $T^{-1} : \operatorname{ran} T \to \mathcal{X}$ is bounded; thus

$$\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\|\,\|Tx\|$$

and $T$ is bounded below. This implies that 0 is not an approximate eigenvalue, and so $0 \notin \sigma_{ap}(T)$. □

The approximate point spectrum is never empty:

***9.5.5. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X})$. Then $\partial\sigma(T) \subset \sigma_{ap}(T)$.*

*Proof.* Let $\lambda \in \partial\sigma(T)$. As $\sigma(T)$ is closed, we have $\lambda \in \sigma(T)$. Let $\{\lambda_n\} \subset \mathbb{C} \setminus \sigma(T)$ with $\lambda_n \to \lambda$. If $\|(T - \lambda_n I)^{-1}\| \leq c$ for all $n$, choose $n$ such that $|\lambda - \lambda_n| < 1/c$; then

$$\|(T - \lambda I) - (T - \lambda_n I)\| = |\lambda - \lambda_n| < \frac{1}{c} \leq \|(T - \lambda_n I)^{-1}\|^{-1}.$$

By Proposition 6.2.3, we obtain that $T - \lambda I \in GL(\mathcal{X})$, a contradiction. So we have that $\|(T - \lambda_n I)^{-1}\| \to \infty$. Now choose, for each $n$, an element $y_n \in \mathcal{X}$ such that $\|y_n\| = 1$ and $\|(T - \lambda_n I)^{-1} y_n\| > \|(T - \lambda_n I)^{-1}\| - 1/n$. Put $x_n = \frac{(T-\lambda_n I)^{-1} y_n}{\|(T-\lambda_n I)^{-1} y_n\|}$. Then $\|x_n\| = 1$ for all $n$, and

$$\begin{aligned}
\|(T - \lambda I)x_n\| &= \|(T - \lambda_n I)x_n + (\lambda_n - \lambda)x_n\| \\
&= \left\| \frac{y_n}{\|(T - \lambda_n I)^{-1} y_n\|} + (\lambda_n - \lambda)x_n \right\| \\
&\leq \frac{1}{\|(T - \lambda_n I)^{-1}\| - 1/n} + |\lambda_n - \lambda| \xrightarrow[n\to\infty]{} 0,
\end{aligned}$$

showing $\lambda \in \sigma_{ap}(T)$. □

Now let us calculate some spectra. We know that if $\mathcal{X}$ is finite-dimensional we may see any $T \in \mathcal{B}(\mathcal{X})$ as a matrix, and $\sigma(T)$ are precisely the eigenvalues. So we will concentrate on infinite-dimensional $\mathcal{X}$. Finding the spectrum (and its parts) for an operator involves very ad-hoc computations, is often hard, and maybe impossible; there is no general recipe.

*9.5.6. Example.* Let $\mathcal{X} = \ell^p(\mathbb{N})$, $1 \leq p \leq \infty$. Fix $b \in \ell^\infty(\mathbb{N})$. Let $M_b \in \mathcal{B}(\ell^p(\mathbb{N}))$ be the multiplication operator given by

$$(M_b x)(n) = b(n)x(n).$$

Then $\|M_b\| = \|b\|_\infty$, and

$$\begin{aligned}
\sigma(M_b) &= \overline{\{b(n) : \ n\}} \\
\sigma_p(M_b) &= \{b(n) : \ n\} \\
\sigma_r(M_b) &= \varnothing \\
\sigma_c(M_b) &= \sigma(T) \setminus \{\{b(n) : \ n\} \\
\sigma_{ap}(M_b) &= \overline{\{b(n) : \ n\}}
\end{aligned}$$

(Exercise 9.5.8).

*9.5.7. Example*. Let $\mathcal{X} = C[0,1]$ with $\|f\|_\infty = \max\{|f(t)| : \ t \in [0,1]\}$, fix $g \in \mathcal{X}$ and let $T \in \mathcal{B}(C[0,1])$ be given by

$$(Tf)(x) = g(x)f(x).$$

Then $\sigma(T) = g([0,1])$. To begin proving this, note that $g([0,1])$ is compact, since $g$ is continuous. For $\lambda \notin g([0,1])$, we have that $(Sf)(x) = \frac{1}{g(x)-\lambda} f(x)$ is an inverse for $T - \lambda I$ (the quotient offers no issue since $g(x) - \lambda$ is bounded away from 0 by the compactness of $g([0,1])$). So $\sigma(T) \subset g([0,1])$. For any $\lambda \in g([0,1])$, if we write $\lambda = g(x_0)$ then we have that $h(x) = (g(x)-\lambda)f(x)$ satisfies $h(x_0) = 0$. Since

$$\|1 - h\|_\infty \geq |1 - h(x_0)| = |1 - 0| = 1,$$

$\operatorname{ran}(T - \lambda I)$ cannot be dense, and so $\sigma(T) = g([0,1])$. Regarding eigenvalues, if $Tf = \lambda f$, then

$$(g(x) - \lambda)\, f(x) = 0, \qquad x \in [0,1].$$

This requires $g(x) = \lambda$ for all those $x$ with $f(x) \neq 0$. As the set $V = \{|f| > 0\}$ is open by continuity, we have that $g|_V = \lambda$. So $\lambda \in \sigma_p(T)$ if and only if the set $\{g = \lambda\}$ has nonempty interior. When $\lambda$ is not an eigenvalue, we showed above that $\operatorname{ran}(T - \lambda I)$ cannot be dense, so $\sigma_c(T) = \varnothing$. As for the approximate point spectrum, given $\lambda \in g([0,1])$ there exists $x_0 \in [0,1]$ with $g(x_0) = \lambda$. For each $n$, by continuity there exists $\delta_n > 0$ such that $|g(x) - \lambda| < \frac{1}{n}$ on $(x_0 - \delta_n, x_0 + \delta_n)$. Let $f_n \in C[0,1]$ with $f_n(x_0) = 1$, $|f_n(x)| \leq 1$ for all $x$, and $f_n = 0$ outside of $(x_0 - \delta_n, x_0 + \delta_n)$. Then $\|f\|_\infty = 1$ and $|(g(x) - \lambda)f_n(x)| \leq \frac{1}{n}$ for all $x$. That is, $\|(T - \lambda)f_n\|_\infty \to 0$, with $\|f_n\|_\infty = 1$ for all $n$. Thus $\sigma_{ap}(T) = g([0,1])$.

We have found

$$\sigma(T) = g([0,1]), \quad \sigma_p(T) = \{\lambda \in g([0,1]) : \ \{g = \lambda\} \ \text{ has nonempty interior}\},$$

$$\sigma_c(T) = \varnothing, \qquad \sigma_r(T) = g([0,1]) \setminus \sigma_p(T) \qquad \sigma_{ap}(T) = g([0,1]).$$

Let us consider now the "same" example but on the space $L^2[0,1]$.

*9.5.8. Example*. Let $\mathcal{X} = L^2[0,1]$ with the usual Hilbert space norm coming from Lebesgue measure, $\|f\|_2 = \left(\int_{[0,1]} |f|^2\right)^{1/2}$, and $T \in \mathcal{B}(L^2[0,1])$ given by $Tf = gf$ for a fixed $g \in L^\infty[0,1]$. Let us consider the eigenvalues first. If $Tf = \lambda f$, this means that $g(x)f(x) = \lambda f(x)$. So on the set $Z = \{f \neq 0\}$ we get $g = \lambda$. If $f \neq 0$ in $L^2$ this means that $m(Z) > 0$. Hence $g\,1_Z = \lambda\,1_Z$. It follows that a necessary and sufficient condition for $\lambda \in \sigma_p(T)$ is that the set $\{g = \lambda\}$ has positive measure.

If $\lambda \notin \operatorname{ess\,ran} g$, then there exists $\delta > 0$ such that $|\lambda - g(x)| > \delta$ a.e. In such case $\frac{1}{g-\lambda} \in L^\infty[0,1]$ and the multiplication operator by $\frac{1}{g-\lambda}$ is an inverse for $T - \lambda I$; that shows that $\sigma(T) \subset \operatorname{ess\,ran} g$. Conversely, if $\lambda \in \operatorname{ess\,ran} g$ then for each $\varepsilon > 0$ we have that $m(g^{-1}(B_\varepsilon(\lambda))) > 0\}$. Let $E_n = g^{-1}(B_{1/n}(\lambda))$. Then $m(E_n) > 0$ for all $n$. Let $f_n = m(E_n)^{-1/2}\, 1_{E_n}$. Then $\|f_n\|_2 = 1$ and

$$\|(T-\lambda I)f_n\|_2^2 = m(E_n)^{-1} \int_{[0,1]} |g-\lambda|^2\, 1_{E_n} \le \frac{1}{n^2}\, m(E_n)^{-1}\, m(E_n) = \frac{1}{n^2}.$$

Hence $\lambda \in \sigma_{ap}(T) \subset \sigma(T)$. This shows that $\operatorname{ess\,ran} g \subset \sigma_{ap}(T) \subset \sigma(T)$, and so $\sigma(T) = \sigma_{ap}(T) = \operatorname{ess\,ran} g$. If $\lambda \in \sigma_r(T)$, then $\lambda \notin \sigma_p(T)$ and $\overline{\operatorname{ran}}(T-\lambda I)$ is a proper subspace of $\mathcal{X}$. Then there exists nonzero $h \in (\operatorname{ran}(T-\lambda I))^\perp$. That is,

$$0 = \langle (T-\lambda I)f, h\rangle = \int_{[0,1]} (gf - \lambda f)\bar{h} = \int_{[0,1]} f\, \overline{(\bar{g}-\bar{\lambda})h} = \langle f, (\bar{g}-\bar{\lambda})h\rangle,$$

for every $f \in L^2[0,1]$. This implies that $g = \lambda$ on the set $\{|h| > 0\}$, which has positive measure if $h \neq 0$. But then $\lambda$ is an eigenvalue, a contradiction. Thus $\sigma_r(T) = \varnothing$. So we have

$$\sigma_p(T) = \{\lambda : \ m(\{g = \lambda\}) > 0\}, \qquad \sigma_r(T) = \varnothing,$$
$$\sigma_c(T) = \operatorname{ess\,ran} g \setminus \sigma_p(T), \qquad \sigma_{ap}(T) = \operatorname{ess\,ran} g,$$

where equality for $\sigma_c(T)$ follows from (9.17).

*9.5.9. Example*. Let $\mathcal{X} = C[0,1]$ with $\|f\|_\infty = \max\{|f(t)| : \ t \in [0,1]\}$. Let $T \in \mathcal{B}(\mathcal{X})$ be given by

$$(Tf)(x) = xf(x) + \int_0^x f(t)\,dt.$$

It is a useful observation that we can write

$$(T-\lambda I)f(x) = \frac{d}{dx}\Big[(x-\lambda)\int_0^x f(t)\,dt\Big]. \tag{9.24}$$

If $\lambda \in \mathbb{C} \setminus [0,1]$, we can simply check that

$$(Sg)(x) = \frac{g(x)}{x-\lambda} - \frac{1}{(x-\lambda)^2}\int_0^x g(t)\,dt$$

(easily deduced from (9.24)) is an inverse for $T - \lambda I$ (the boundedness of $S$ is automatic from the Inverse Mapping Theorem). Thus $\sigma(T) \subset [0,1]$. We will confirm that this is an equality when we calculate the residual spectrum below.

The expression in (9.24) allows us to immediately see that $T-\lambda I$ is injective for all $\lambda$. For if $(T-\lambda I)f = 0$, then there exists $c \in \mathbb{C}$ such that

$$(x-\lambda)\int_0^x f(t)\,dt = c, \qquad x \in [0,1].$$

Evaluating at $x=\lambda$, we get that $c=0$. But then $\int_0^x f = 0$ for all $x$ (other than $\lambda$, but then also $\lambda$ by continuity); differentiating, $f=0$. So $T-\lambda I$ is injective when $\lambda \in [0,1]$. In particular, $\sigma_p(T)=\varnothing$.

Let us now focus on the range of $T-\lambda I$. Fix $\lambda \in (0,1]$. If $g \in \operatorname{ran}(T-\lambda I)$, then

$$g(x) = (x-\lambda)f(x) + \int_0^x f = \frac{d}{dx}\Big[(x-\lambda)\int_0^x f(t)\,dt\Big].$$

for some $f \in C[0,1]$. Suppose that there exists $g \in \operatorname{ran} T-\lambda I$ such that $\|1-g\|_\infty < 1$. We also have $\|1-\operatorname{Re} g\|_\infty < 1$. If $g=(T-\lambda I)f$, the linearity and continuity of the real-part function allows us to write $\operatorname{Re} g = (T-\lambda I)(\operatorname{Re} f)$. So we may assume without loss of generality that $f$ and $g$ are real-valued. From $\|1-g\|_\infty < 1$, there exists $\varepsilon \in (0,1)$ such that $1-\varepsilon \le g(x) \le 1+\varepsilon$ for all $x \in [0,1]$ (just take $\varepsilon = \|1-g\|_\infty$). Using (9.24) and integrating between $\lambda$ and $x$ for $x<\lambda$ (so that integrating reverses inequalities),

$$(1+\varepsilon)(x-\lambda) \le (x-\lambda)\int_0^x f \le (1-\varepsilon)(x-\lambda).$$

For any $x<\lambda$ we can divide by $x-\lambda$, the inequalities are reversed again, and so

$$1-\varepsilon \le \int_0^x f \le 1+\varepsilon, \qquad 0<x<\lambda. \tag{9.25}$$

Now taking $x$ sufficiently close to 0 leads to a contradiction. It follows that when $\lambda \in (0,1]$ we have $\|1-g\|_\infty \ge 1$ for all $g \in \operatorname{ran} T-\lambda I$, and thus $\sigma_r(T) \supset (0,1]$. When $\lambda=0$ from (9.24) we obtain $g(0)=0$, and so $\|1-g\|_\infty \ge |1-0| = 1$. Then $\sigma_r(T)=[0,1]$. We immediately get $\sigma_c(T)=\varnothing$ and $\sigma(T)=[0,1]$. Finally, for the approximate point spectrum, since $\sigma(T)=[0,1]\subset\mathbb{C}$, every point is a boundary point. Then $\lambda\in\sigma_{ap}(T)$ by Proposition 9.5.5. Thus

$$\sigma(T)=\sigma_r(T)=\sigma_{ap}(T)=[0,1], \qquad \sigma_p(T)=\sigma_c(T)=\varnothing.$$

*9.5.10. Example.* The **unilateral shift** $S$ is a common example of an operator that arises very naturally and has non-trivial properties. Fix $p$, with $1<p\le\infty$, and let $q$ be its conjugate, such that $\frac{1}{p}+\frac{1}{q}=1$. Define $S:\ell^p(\mathbb{N})\to\ell^p(\mathbb{N})$ by

$$Sx = (0, x_1, x_2, \ldots). \tag{9.26}$$

It is clear that $\|Sx\|_p = \|x\|_p$ for all $x$, so $S$ is an isometry; in particular, it is bounded with $\|S\|=1$. Consider also $T:\ell^q(\mathbb{N})\to\ell^q(\mathbb{N})$ given by

$$Tx = (x_2, x_3, \ldots).$$

Then, when $q < \infty$,

$$\|Tx\|_q^q = \sum_n |x_{n+1}|^q \leq \sum_n |x_n|^q = \|x\|_q^q.$$

And if $q = \infty$ we also have $\|Tx\|_q \leq \|x\|_q$. So $\|T\| \leq 1$, and $T$ is also bounded; it is easy to check that $\|T\| = 1$ (for instance, $\|Te_2\| = \|e_1\| = 1$). Note that if we consider $T$ acting on $\ell^p(\mathbb{N})$, we have $TSx = x$ for all $x$, so $TS = I$ in $\mathcal{B}(\ell^p(\mathbb{N}))$. But

$$STx = (0, x_2, x_3, \ldots),$$

so $ST \neq I$.

From $\|S\| = 1$ we get (via Theorem 9.2.16) that $\operatorname{spr}(S) \leq 1$, so $\sigma(S) \subset \overline{\mathbb{D}}$. Similarly, $\sigma(T) \subset \overline{\mathbb{D}}$. Since $(\ell^p(\mathbb{N}))^* = \ell^q(\mathbb{N})$ (Proposition 5.6.3), $S^* \in \mathcal{B}(\ell^q(\mathbb{N}))$. We have, for $y \in \ell^q(\mathbb{N})$, $x \in \ell^p(\mathbb{N})$,

$$(S^*y)x = y(Sx) = \sum_{n\geq 2} y_n x_{n-1} = \sum_{n \geq 1} y_{n+1} x_n = (Ty)\, x.$$

So $S^* = T$ and $\sigma(T) = \sigma(S)$ by Proposition 9.5.1. The same computation gives us

$$(T^*x)y = x(Ty) = \sum_{n\geq 2} x_{n-1} y_n = (Sx)y,$$

so $T^* = S$.

Now let us look at the spectrum. If $\lambda \in \sigma_p(S)$, then there exists nonzero $x$ with $Sx = \lambda x$. This gives us the equations

$$\lambda x_1 = 0,\ \lambda x_2 = x_1,\ \lambda x_3 = x_2,\ \ldots$$

We have $\lambda \neq 0$, since $\|\lambda x\|_p = \|Sx\|_p = \|x\|_p \neq 0$. But the the equalities above give successively $x_1 = 0$, $x_2 = 0$, etc. So $x = 0$ and $\lambda$ cannot be an eigenvalue. That is, $\sigma_p(S) = \varnothing$. For $T$, on the other hand, let $\lambda \in \mathbb{D}$. Then $(\lambda, \lambda^2, \lambda^2, \ldots) \in \ell^q(\mathbb{N})$, and

$$T(\lambda, \lambda^2, \lambda^3, \ldots) = (\lambda^2, \lambda^3, \lambda^4, \ldots) = \lambda\, (\lambda, \lambda^2, \lambda^3, \ldots).$$

So $\lambda \in \sigma_p(T)$, and then $\mathbb{D} \subset \sigma_p(T)$. Moreover, if $|\lambda| = 1$ then $\lambda \notin \sigma_p(T)$: indeed, if $Tx = \lambda x$, then we get equalities

$$x_2 = \lambda x_1,\ x_3 = \lambda x_2, \ldots$$

These equalities require the eigenvectors for $\lambda$ to be of the form $x_1\,(1, \lambda, \lambda^2, \ldots)$, which is not in $\ell^p(\mathbb{N})$ since $|\lambda| = 1$. Thus $\sigma_p(T) = \mathbb{D}$.

From $\mathbb{D} \subset \sigma(T) \subset \overline{\mathbb{D}}$ and the fact that the spectrum is closed, we get $\sigma(T) = \overline{\mathbb{D}}$ and then $\sigma(S) = \overline{\mathbb{D}}$.

For any $\lambda \in \mathbb{D}$ we have $\dim\ker(T - \lambda I) = 1$ (Exercise 9.5.14), and also

$$\|(S - \lambda I)x\|_p = \|Sx - \lambda x\|_p \geq \|Sx\|_p - |\lambda|\, \|x\|_p = (1 - |\lambda|)\|x\|_p, \tag{9.27}$$

so $S - \lambda I$ is bounded below; by Proposition 9.5.4, $\lambda \notin \sigma_{ap}(S)$. Thus $\sigma_{ap}(S) \subset \mathbb{T} = \partial\sigma(S)$. From Proposition 9.5.5 we have that $\partial\sigma(S) \subset \sigma_{ap}(S)$, and so $\sigma_{ap}(S) = \partial\sigma(S) = \mathbb{T}$. As any $\lambda \in \mathbb{D}$ is not an approximate eigenvalue and

$\sigma(S) = \sigma_{ap}(S) \cup \sigma_r(S)$, we get $\sigma_r(S) \supset \mathbb{D}$. We also have $\mathbb{T} = \partial\sigma(T) \subset \sigma_{ap}(T)$, so together with $\mathbb{D} = \sigma_p(T) \subset \sigma_{ap}(T)$ we have $\overline{\mathbb{D}} \subset \sigma_{ap}(T)$ and then $\sigma_{ap}(T) = \overline{\mathbb{D}}$.

If $\lambda \in \mathbb{T}$, we showed above that $\lambda \notin \sigma_p(T)$; so $T - \lambda I$ is injective. By Proposition 9.4.6, $\operatorname{ran}(S - \lambda I)$ is dense, and so $\lambda \in \sigma_c(S)$. That is, $\mathbb{T} \subset \sigma_c(S)$. As $\sigma_c(S) \cap \sigma_r(S) = \varnothing$, we obtain that $\sigma_c(S) = \mathbb{T}$, $\sigma_r(S) = \mathbb{D}$.

Note also that, from Proposition 9.5.3, $\sigma_r(T) \subset \sigma_p(S) = \varnothing$. By (9.17), $\sigma(T) = \sigma_p(T) \cup \sigma_c(T)$, a disjoint union. As we saw above that $\sigma_p(T) = \mathbb{D}$, we obtain that $\sigma_c(T) = \mathbb{T}$.

In summary,

$$\begin{aligned} \sigma_p(S) &= \varnothing \\ \sigma_{ap}(S) &= \mathbb{T} \\ \sigma_r(S) &= \mathbb{D} \\ \sigma_c(S) &= \mathbb{T} \end{aligned} \qquad \begin{aligned} \sigma_p(T) &= \mathbb{D} \\ \sigma_{ap}(T) &= \overline{\mathbb{D}} \\ \sigma_r(T) &= \varnothing \\ \sigma_c(T) &= \mathbb{T} \end{aligned}$$

*9.5.11. Example.* When $p = 1$, calculating the spectrum of the shifts $S$ and $T$ above requires a slight tweak in the strategies.

Consider $S : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$ and $T : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$. We still have that $T = S^*$. The main difference is that for $T : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$, each $\lambda \in \overline{\mathbb{D}}$ is an eigenvalue. That is, $\sigma_p(T) = \overline{\mathbb{D}}$, and so $\sigma_r(T) = \sigma_c(T) = \varnothing$. By Proposition 9.5.3, and since $\sigma_p(S) = \varnothing$,

$$\overline{\mathbb{D}} = \sigma_p(T) = \sigma_p(S^*) \subset \sigma_r(S) \cup \sigma_p(S) = \sigma_r(S).$$

Then, as we did in (9.27), $\sigma_{ap}(S) = \mathbb{T}$. We have

$$\begin{aligned} \sigma_p(S) &= \varnothing \\ \sigma_{ap}(S) &= \mathbb{T} \\ \sigma_r(S) &= \overline{\mathbb{D}}\text{'} \\ \sigma_c(S) &= \varnothing \end{aligned} \qquad \begin{aligned} \sigma_p(T) &= \overline{\mathbb{D}} \\ \sigma_{ap}(T) &= \overline{\mathbb{D}} \\ \sigma_r(T) &= \varnothing \\ \sigma_c(T) &= \varnothing \end{aligned}$$

*9.5.12. Example* If we now consider $S : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ and $T : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$, we have $S = T^*$ but $T$ is not the adjoint of $S$. Everything from Example 9.5.10 works with the exception of the use of Proposition 9.5.3. But we still use (9.27) to conclude that $\mathbb{D} \subset \sigma_r(S)$ and $\sigma_c(S) \subset \mathbb{T}$. Suppose that $y \in \operatorname{ran} S - I$ and $\|1 - y\|_\infty < 1$. Let $\varepsilon = \frac{1}{2}\|1 - y\|_\infty$, and $x \in \ell^\infty(\mathbb{N})$ such that $y = Sx - x$. From $|1 - y_k| < \frac{1}{2}$ we have $y_k = 1 + z_k$, where $|z_k| < \frac{1}{2}$. From $y = Sx - x$ we get

$$y = (-x_1, x_1 - x_2, x_2 - x_3, \ldots).$$

Then $x_1 = -y_1$, $x_2 = x_1 - y_2 = -y_1 - y_2$, $x_k = -y_1 - \cdots - y_k$. This gives us

$$|x_k| = |y_1 + \cdots + y_k| = |k + z_1 + \cdots + z_k| \geq k - \sum_{j=1}^{k} |z_k| \geq k - \frac{1}{2}\,k = \frac{3}{2}\,k.$$

Then $x \notin \ell^\infty(\mathbb{N})$, a contradiction. It follows that $\|1 - y\|_\infty \geq 1$, and hence $\operatorname{ran} T - I$ is not dense; that is, $1 \in \sigma_r(S)$. Now we use a trick. Given $\mu \in \mathbb{T}$, we define $Z : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ by

$$Zx = (\mu x_1, \mu^2 x_2, \mu^3 x_3, \ldots).$$

Then $Z$ is a linear isometric bijection. We have $ZSZ^{-1} = \mu S$. As we also have

$$S - \mu^{-1}\lambda I = \mu^{-1}(\mu S - \lambda I) = \mu^{-1}(ZSZ^{-1} - \lambda ZZ^{-1}) = \mu^{-1} Z(S - \lambda I)Z^{-1}.$$

As $Z$, $Z^{-1}$, and multiplication by $\mu^{-1}$ are isometries, the kernel and the closedness and density or not of the rank are preserved. Thus for any $\mu \in \mathbb{T}$ we have

$$\sigma(ZSZ^{-1}) = \mu\sigma(S), \qquad \sigma_p(ZSZ^{-1}) = \mu\sigma_p(S),$$
$$\sigma_r(ZSZ^{-1}) = \mu\sigma_r(S), \qquad \sigma_c(ZSZ^{-1}) = \mu\sigma_c(S).$$

This allows us to conclude, knowing that $1 \in \sigma_r(S)$ and by running $\mu$ over all of the unit circle, that $\sigma_r(S) \supset \mathbb{T}$. Hence

$$\begin{array}{ll} \sigma_p(S) = \varnothing & \sigma_p(T) = \mathbb{D} \\ \sigma_{ap}(S) = \mathbb{T} & \sigma_{ap}(T) = \overline{\mathbb{D}} \\ \sigma_r(S) = \overline{\mathbb{D}} & \sigma_r(T) = \varnothing \\ \sigma_c(S) = \varnothing & \sigma_c(T) = \mathbb{T} \end{array}$$

The values for $T$ are obtained, for $\sigma_p(T)$ and $\sigma_{ap}(T)$, as before. Since $S = T^*$, Proposition 9.5.3 gives us $\sigma_r(T) \subset \sigma_p(S) = \varnothing$. As $\overline{\mathbb{D}} = \sigma(T) = \sigma_p(T) \cup \sigma_r(T) \cup \sigma_c(T)$ as a disjoint union and $\sigma_p(T) = \mathbb{D}$, we obtain $\sigma_c(T) = \mathbb{T}$. This gives us an example that (9.23) can fail when $\mathcal{X}$ is not reflexive.

*9.5.13. Example.* Let us see that the inclusion (9.19) can be strict. Consider the left unilateral shift $T$, as in Example 9.5.10, but considered as an operator $T : c_0 \to c_0$. The exact same argument as in the example shows that $\sigma_p(T) = \mathbb{D}$. Meanwhile, $T^{**}$ is also the unilateral shift, but now considered as a map $\ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N})$. Now we can also consider eigenvectors $(\lambda, \lambda^2, \ldots)$ with $|\lambda| = 1$, and so $\sigma_p(T^{**}) = \overline{\mathbb{D}}$.

*9.5.14. Example*. Here is a much simpler example. Let $E \in \mathcal{B}(\mathcal{X})$ be an idempotent, i.e. $E^2 = E$. We have $E = 0$ and $E = I_{\mathcal{X}}$ as trivial examples. When $E$ is not trivial, then $\sigma(E) = \{0, 1\}$ and both are eigenvalues. Indeed, if $\lambda \notin \{0, 1\}$, it can be checked directly that

$$(\lambda - E)^{-1} = \frac{1}{\lambda}\,(I_{\mathcal{X}} - E) - \frac{1}{1-\lambda}\,E. \tag{9.28}$$

So $\sigma(E) \subset \{0, 1\}$. Since $E \neq I_{\mathcal{X}}$ we have that $E\mathcal{X}$ is a proper subspace of $\mathcal{X}$, and any element of $E\mathcal{X}$ is an eigenvector for 1. We also have that $(I_x - E)\mathcal{X}$ is a proper subspace and any element of $(I_x - E)\mathcal{X}$ is an eigenvector for 0 (that is, $(I_{\mathcal{X}} - E)\mathcal{X} = \ker E$).

It seems that (9.28) came out of nowhere. But one can directly check that the unital Banach algebra $\mathcal{A}$ generated by $E$ in $\mathcal{B}(\mathcal{X})$ consists precisely of the elements $\alpha E + \beta(I_{\mathcal{X}} - E)$ for $\alpha, \beta \in \mathbb{C}$, so one would expect to find an inverse for $\lambda I_{\mathcal{X}} - E$ in $\mathcal{A}$.

**9.5.15. Proposition.**

*Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ an isometry. Then*

*(i) if $T \in GL(\mathcal{X})$, then $\sigma(T) \subset \mathbb{T}$;*

*(ii) if $T \notin GL(\mathcal{X})$, then $\sigma(T) = \overline{\mathbb{D}}$ and $\sigma_{ap}(T) = \mathbb{T}$.*

*Proof.* Because $\|Tx\| = \|x\|$ for all $x$, we have $\|T\| = 1$. Then $\sigma(T) \subset \overline{\mathbb{D}}$. When $T$ is invertible, $\lambda \in \sigma(T)$ if and only if $\lambda^{-1} \in \sigma(T)$ (Proposition 9.2.9). So $\sigma(T) \subset \mathbb{T}$, as $|\lambda| \leq 1$ and $|\lambda^{-1}| \leq 1$.

Now assume that $T$ is not invertible. As we did with the unilateral shift,

$$\|(T - \lambda I)x\| \geq \|Tx\| - \|\lambda x\| = (1 - |\lambda|)\|x\|. \tag{9.29}$$

So $|\lambda| < 1$ implies that $T - \lambda I$ is bounded below, and so $\sigma_{ap}(T) \subset \mathbb{T}$.

Since $T$ is not invertible, $0 \in \sigma(T)$. By Proposition 9.5.5 $\partial\sigma(T) \subset \sigma_{ap}(T) \subset \mathbb{T}$. So 0 is an interior point in $\sigma(T)$; thus the interior of $\sigma(T)$ is a nonempty open subset of $\mathbb{D}$ with boundary in $\mathbb{T}$. This forces the interior of $\sigma(T)$ to be $\mathbb{D}$. Thus $\sigma(T) = \overline{\mathbb{D}}$. Also, $\mathbb{T} = \partial\sigma(T) \subset \sigma_{ap}(T) \subset \mathbb{T}$, so $\sigma_{ap}(T) = \mathbb{T}$. □

*9.5.16. Example.* The **bilateral shift** $V$ is another natural example. The idea is the same as in the case of the unilateral shift, but we change the index set. We define $V \in \mathcal{B}(\ell^p(\mathbb{Z}))$, $p \in (1, \infty)$, in the following way. As usual, denote the canonical basis by $\{e_n\}_{n\in\mathbb{Z}}$. Then we let $V$ to be the linear operator induced by

$Ve_n = e_{n+1}$. We have, for any $x \in \operatorname{span}\{e_n:\ n\}$,

$$\|Vx\|_p = \Big\| \sum_n x_n e_{n+1} \Big\|_p = \Big( \sum_n |x_n|^p \Big)^{1/p} = \|x\|_p$$

(for $p = \infty$ the norm is different, but we still have $\|Vx\| = \|x\|$). So $V$ is isometric on a dense subspace and extends to the whole space as an isometry. Alternatively, we can define $V$ directly in an analog way to (9.26), but the notation is awkward with the sequence being infinite on both sides.

If we define the **right bilateral shift** $W$ by $We_n = e_{n-1}$, the same arguments as above give us that $W$ is an isometry, and $WV = VW = I$. As $V$ is an invertible isometry, we know that $\sigma(V) \subset \mathbb{T}$ by Proposition 9.5.15. Similarly, $\sigma(W) \subset \mathbb{T}$. Now, for $\lambda \in \mathbb{T}$,

$$V - \lambda I = V - \lambda VW = -\lambda V\,(W - \bar{\lambda} I),$$

so $\lambda \in \sigma(V) \iff \bar{\lambda} \in \sigma(W)$. We will show that $\sigma(V) = \sigma(W) = \mathbb{T}$ by adapting the trick from Example 9.5.12. Fix $\mu \in \mathbb{T}$, and define $Z \in \mathcal{B}(\ell^p(\mathbb{Z}))$ by

$$Ze_n = \mu^n e_n$$

and extend by linearity. Because $Z$ multiplies each entry by a scalar of absolute value 1, it is an isometry. It is clear that $Z^{-1}e_n = \mu^{-n}e_n$. Then

$$ZVZ^{-1}e_n = \mu^{-n} Ze_{n+1} = \mu\, e_{n+1} = \mu\, Ve_n.$$

Thus $ZVZ^{-1} = \mu V$. It follows, as $V - \lambda I$ is invertible if and only if $Z(V - \lambda I)Z^{-1}$ is invertible, that

$$\sigma(V) = \sigma(ZVZ^{-1}) = \sigma(\mu V) = \mu\sigma(V).$$

Thus $\lambda \in \sigma(V) \iff \mu\lambda \in \sigma(V)$, and this can be done for all $\mu \in \mathbb{T}$. We conclude that $\sigma(V) = \sigma(W) = \mathbb{T}$. There is no qualitative difference between $V$ and $W$, as their distinction lies on a reordering of the canonical basis, so we will only think about $\sigma(V)$.

We can consider $V^* : \ell^q(\mathbb{Z}) \to \ell^q(\mathbb{Z})$. We have, for $y \in \ell^q(\mathbb{Z})$, $x \in \ell^p(\mathbb{Z})$,

$$(V^*y)x = y(Vx) = \sum_{n\in\mathbb{Z}} y_n x_{n-1} = \sum_{n\in\mathbb{Z}} y_{n+1} x_n = (Wy)\,x.$$

So $V^* = W$ and similarly $W^* = V$. As with the unilateral shift, a straightforward computation shows that since $p < \infty$ $V$ has no eigenvalues, so $\sigma_p(V) = \varnothing$. For any $\lambda \in \mathbb{T}$, since $\lambda \in \sigma(V)$ we know that $V - \lambda I$ is not invertible. Since there are no eigenvalues, $V - \lambda I$ and $W - \lambda I$ are injective; by Proposition 9.4.6, $\operatorname{ran}(V - \lambda I)$ is dense. The range cannot be closed because that would make $T - \lambda I$ invertible via the Inverse Mapping Theorem (6.3.6). So $\lambda \in \sigma_c(V)$ and $\lambda \in \sigma_{ap}(V)$ by

(9.17). Therefore

$$\begin{aligned} \sigma_p(V) &= \varnothing & \sigma_p(W) &= \varnothing \\ \sigma_{ap}(V) &= \mathbb{T} & \sigma_{ap}(W) &= \mathbb{T} \\ \sigma_r(V) &= \varnothing & \sigma_r(W) &= \varnothing \\ \sigma_c(V) &= \mathbb{T} & \sigma_c(W) &= \mathbb{T} \end{aligned}$$

*9.5.17. Example.* When we consider the bilateral shift on $\ell^\infty(\mathbb{Z})$, things change a bit. We can still define $V$ and $W$ and they are isometric and inverses of each other, and the same arguments as in Example 9.5.16 show that $\sigma(V) = \sigma(W) = \mathbb{T}$. When we consider the parts of the spectrum things are different, though, as it has already happened with the unilateral shift.

We consider $V \in \mathcal{B}(\ell^1(\mathbb{Z}))$, and then $W = V^* \in \mathcal{B}(\ell^\infty(\mathbb{Z}))$. For $\lambda \in \mathbb{T}$, we can define $x = \sum_n \lambda^n e_n \in \ell^\infty(\mathbb{Z})$, with $\|x\| = 1$, and we have that $Wx = \lambda x$. So $\sigma_p(W) = \mathbb{T}$. We immediately get by definition that $\sigma_r(W) = \sigma_c(W) = \varnothing$, and $\sigma_{ap}(W) = \mathbb{T}$. From $\sigma_p(V) = \varnothing$, $\sigma_p(W) = \mathbb{T}$ and (9.20), we have $\sigma_r(V) = \mathbb{T}$. This forces $\sigma_c(V) = \varnothing$. From Proposition 9.5.5, $\mathbb{T} = \partial\sigma(V) \subset \sigma_{ap}(V)$. Thus, for $V \in \mathcal{B}(\ell^1(\mathbb{N}))$ and $W \in \mathcal{B}(\ell^\infty(\mathbb{N}))$,

$$\begin{aligned} \sigma_p(V) &= \varnothing & \sigma_p(W) &= \mathbb{T} \\ \sigma_{ap}(V) &= \mathbb{T} & \sigma_{ap}(W) &= \mathbb{T} \\ \sigma_r(V) &= \varnothing & \sigma_r(W) &= \varnothing \\ \sigma_c(V) &= \mathbb{T} & \sigma_c(W) &= \varnothing \end{aligned}$$

We saw above that the right shift on $\ell^p(\mathbb{N})$ has a disk of eigenvalues; such property can be seen as a particular case of the following:

***9.5.18. Proposition.***

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$. If $T$ is not invertible and there exist $c > 0$ and a function $R : \mathcal{X} \to \mathcal{X}$ (not necessarily linear) such that $\|Rx\| \leq c\|x\|$ for all $x \in \mathcal{X}$ and $TR = I$, then $B_{1/2c}(0) \subset \sigma_p(T)$.*

*Proof.* From $TR = I$ we have that $T$ is surjective. By the Inverse Mapping Theorem (6.3.6), $T$ is not injective; so $\ker T \neq \{0\}$. Fix $\lambda \in B_{1/2c}(0)$ and

$x_0 \in \ker T$ with $\|x_0\| = 1$. Define

$$x = \sum_{k=0}^{\infty} \lambda^k R^k x_0.$$

This converges because $\|\lambda^k R^k x_0\| \le |\lambda| c < 1$ and so $\|x\| \le \sum_k |\lambda|^k c^k < \infty$. Also, $x \neq 0$ because the term for $k = 0$ satisfies $\|x_0\| = 1$, and for the rest of the series,

$$\left\| \sum_{k=1}^{\infty} \lambda^k R^k x_0 \right\| \le \sum_{k=1}^{\infty} |\lambda|^k c^k < \sum_{k=1}^{\infty} 2^{-k} = 1$$

Thus

$$\|x\| \ge \|x_0\| - \left\| \sum_{k=1}^{\infty} \lambda^k R^k x_0 \right\| > 1 - \sum_{k=1}^{\infty} |\lambda|^k c^k > 0.$$

Finally, since $T$ is bounded and $Tx_0 = 0$,

$$Tx = \sum_{k=1}^{\infty} \lambda^k T R^k x_0 = \sum_{k=1}^{\infty} \lambda^k R^{k-1} x_0 = \lambda \sum_{k=1}^{\infty} \lambda^{k-1} R^{k-1} x_0 = \lambda x. \qquad \square$$

The argument in Proposition 9.5.18 works for $\lambda \in B_{1/c}(0)$, so in principle one gets $B_{1/c}(0) \subset \sigma_p(T)$. That is the case for the left unilateral shift, for instance. The problem is that in the general there does not seem to be an obvious way to guarantee that the eigenvector is nonzero without the restriction $|\lambda| c \le \frac{1}{2}$.

Here is a more friendly way to state a condition that guarantees tons of eigenvalues.

***9.5.19. Corollary.***

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$. If $T$ is surjective but not injective, then there exists $r > 0$ such that $B_r(0) \subset \sigma_p(T)$.*

*Proof.* Let $\mathcal{Y} = \ker T$. The map $\tilde{T} : \mathcal{X}/\mathcal{Y} \to \mathcal{X}$ given by $\tilde{T}(x + \mathcal{Y}) = Tx$ is bounded, linear, and bijective (Exercise 9.5.19); so invertible by the Inverse Mapping Theorem (6.3.6). That is, there exists $S : \mathcal{X} \to \mathcal{X}/\mathcal{Y}$, bounded, with $\tilde{T}S = I_{\mathcal{X}}$. Fix $\varepsilon > 0$. For each $x \in \mathcal{X}$, choose a representative $Rx$ of $Sx$ with $\|Rx\| \le (1+\varepsilon)\|Sx\|$ ($Rx$ exists by definition of the quotient norm). Then $TRx = \tilde{T}(Rx + \mathcal{Y}) = \tilde{T}Sx = x$. And $\|Rx\| \le (1+\varepsilon)\|Sx\| \le (1+\varepsilon)\|S\|\,\|x\|$. Thus we can apply Proposition 9.5.18 to $T$, with $c = (1+\varepsilon)\|S\|$ and $r = \frac{1}{(1+\varepsilon)\|S\|}$. As $\sigma_p(T)$ contains all balls $B_{\frac{1}{(1+\varepsilon)\|S\|}}(0)$ for all $\varepsilon > 0$, we can have $r = \frac{1}{\|S\|}$. $\square$

The proof above does not contradict Remark 6.2.5, for the right-inverse we constructed here is not necessarily linear.

*9.5.20. Example.* Consider $T : C[0,1] \to C[0,1]$, where $(Tf)(t) = f(t/2)$. Then $T$ is surjective but not injective (Exercise 9.5.20). Corollary 9.5.19 tells us that there exists $r > 0$ such that if $|\lambda| < r$, then $\lambda$ is an eigenvalue for $T$. If one is faced with the problem of finding the eigenfunction for a certain $\lambda$, we are looking at finding $f$ satisfying

$$f(t/2) = \lambda f(t), \qquad t \in [0,1], \tag{9.30}$$

which might not look too enticing. In particular, could we read $r$ from (9.30)? Because nonzero $f$ will not exist for $|\lambda|$ big enough. The problem does not look obvious, but in Proposition 9.5.18 we have a recipe to find the eigenfunctions. Since

$$\ker T = \{f : \ f(t) = 0, \ 0 \leq t \leq \tfrac{1}{2}\},$$

we can define

$$Rf = \begin{cases} f(2t), & 0 \leq t \leq 1/2 \\ f(1), & 1/2 < t \leq 1 \end{cases}$$

Then $Rf \in C[0,1]$, $TRf = f$, and $\|Rf\| = \|f\|$. So $c = 1$, which allows us to take $r = 1$ (being a concrete example, we can check directly whether the constructed eigenvector is nonzero, so the restriction from Proposition 9.5.18 is not needed). Let $g_0(t) = (2t-1)\,1_{(1/2,1]}(t)$. It is not hard to check by induction that

$$(R^k g_0)(t) = \begin{cases} 0, & t \leq 2^{-k-1} \\ 2^{k+1}t - 1, & 2^{-k-1} < t \leq 2^{-k} \\ 1, & 2^{-k} < t \end{cases}$$

This gives us in particular

$$R^{k+1} g_0(t/2) = R^k g_0(t). \tag{9.31}$$

Then, for $\lambda$ with $|\lambda| < 1$, an eigenfunction for $\lambda$ is—after a bit of work to lump together the constant chunks (Exercise 9.5.21)—

$$f_\lambda(t) = \sum_{k=0}^{\infty} \lambda^k\, R^k g_0 = \sum_{k=0}^{\infty} \lambda^k \left(2^{k+1}t - \frac{1-2\lambda}{1-\lambda}\right) 1_{[2^{-k-1},2^{-k}]}(t).$$

Indeed, noting that $g_0\left(\frac{t}{2}\right) = 0$ since $t/2 \leq 1/2$, and using (9.31),

$$\begin{aligned} f_\lambda\left(\tfrac{t}{2}\right) &= \sum_{k=0}^{\infty} \lambda^k\, R^k g_0\left(\tfrac{t}{2}\right) = \sum_{k=1}^{\infty} \lambda^k\, R^k g_0\left(\tfrac{t}{2}\right) = \sum_{k=1}^{\infty} \lambda^k\, R^{k-1} g_0(t) \\ &= \lambda \sum_{k=0}^{\infty} \lambda^k\, R^k g_0(t) = \lambda\, f_\lambda(t). \end{aligned}$$

When $\lambda = 1/2$, the formula for $f_\lambda$ gives the obvious eigenvector $f_{1/2}(t) = 2t$. But for other values of $\lambda$ the eigenfunctions are less obvious; in particular they are not differentiable at the points $t = 2^{-k}$. They are fairly simple, though, as they are piecewise linear; note that while $f_\lambda$ is defined as an infinite sum, the terms are nonzero on distinct intervals. So what we are doing is take a segment of slope 2 on $\left[\frac{1}{2},1\right]$; then slope $4\lambda$ on $\left[\frac{1}{4},\frac{1}{2}\right]$; then $8\lambda^2$ on $\left[\frac{1}{8},\frac{1}{4}\right]$, etc. On each dyadic segment (from right to left) we multiply the slope by $2\lambda$, the 2 accounting for the change of variable $t \longmapsto \frac{t}{2}$ and $\lambda$ for the eigenvalue.

So far we know that $\mathbb{D} \subset \sigma_p(T) \subset \sigma(T) \subset \overline{\mathbb{D}}$, which immediately tells us that $\sigma(T) = \overline{\mathbb{D}}$. Any constant function is an eigenvector for 1, so $1 \in \sigma_p(T)$. For $\lambda \in \mathbb{T}\setminus\{1\}$, from $f(t/2) = \lambda f(t)$ we get that $f(0) = 0$. Also, iterating the equality and using the continuity of $f$ and that $|\lambda^{-k}| = 1$ for all $k$,

$$f(t) = \lambda^{-1} f(t/2) = \lambda^{-2} f(t/4) = \cdots = \lambda^{-k} f(t/2^k) \xrightarrow[k\to\infty]{} 0.$$

Thus $f = 0$ and $\lambda \notin \sigma_p(T)$, which proves that $\sigma_p(T) = \mathbb{D}\cup\{1\}$. For $\lambda \in \mathbb{T}\setminus\{1\}$,

$$(T-\lambda I)f(t) = f(t/2) - \lambda f(t).$$

The dual of $C[0,1]$ is the space of regular Borel measures (Theorem 5.6.15). We have $(T^*\mu)f = \mu(Tf)$. So

$$\begin{aligned}(T^*\mu)(E) &= \int_{[0,1]} 1_E\, d(T^*\mu) = \int_{[0,1]} (T1_E)\, d\mu = \int_{[0,1]} 1_E(t/2)\, d\mu(t)\\ &= \int_{[0,1]} 1_{2(E\cap[0,1/2])}(t)\, d\mu(t) = \mu(2(E\cap[0,1/2])).\end{aligned}$$

If $\mu \in \ker(T^*-\lambda I)$, this means that $\mu(2(E\cap[0,1/2])) - \lambda\mu(E) = 0$ for all Borel $E \subset [0,1]$. Then $\mu(E) = \lambda^{-1}\mu(2(E\cap[0,1/2]))$; iterating this,

$$\mu(E) = \lambda^{-k}\mu(2^k(E\cap[0,2^{-k}])).$$

Using $k$ big enough, we conclude that $\mu([\delta,1]) = 0$ for all $\delta > 0$. Thus

$$\mu((0,1]) = \bigcup_n \mu([1/n,1]) = 0$$

by continuity of the measure. That is, $\mu$ can only be concentrated at zero; if $\mu \neq 0$, then $\mu = c\delta_0$ for some $c \in \mathbb{C}$. But

$$(T^*\delta_0)f = \delta_0(Tf) = f(0/2) = f(0) = \delta_0 f.$$

So $(T^*-\lambda I)\delta_0 = (1-\lambda)\delta_0$; as $\lambda \neq 1$, $\ker(T^*-\lambda I) = \{0\}$. Now Proposition 9.4.6 gives us that $\operatorname{ran}(T-\lambda I)$ is dense, and hence $\lambda \in \sigma_c(T)$. In summary,

$$\sigma(T) = \overline{\mathbb{D}}, \qquad \sigma_p(T) = \mathbb{D}\cup\{1\}, \qquad \sigma_c(T) = \mathbb{T}\setminus\{1\}, \qquad \sigma_r(T) = \varnothing.$$

We record another straightforward application of Proposition 9.5.18.

***9.5.21. Corollary.***

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$. If $0 \in \sigma(T)$ is an isolated point, then $T$ is not surjective.*

*Proof.* If $T$ is injective, it cannot be surjective because then it would be invertible by the Inverse Mapping (Theorem 6.3.6), contradicting that $0 \in \sigma(T)$.

If $T$ is not injective, it cannot be surjective because by Corollary 9.5.19 $\sigma(T)$ would have a ball around 0, contradicting the hypothesis that 0 is isolated. ☐

*9.5.22. Example.* For some examples with a different flavour, fix $p \in (1, \infty]$ and consider the **Hardy operators**, also called **Cesàro operators**, $T_p : L^p[0,1] \to L^p[0,1]$ given by

$$(T_p f)(x) = \frac{1}{x} \int_0^x f.$$

We can similarly consider, with the same formula, $S_p : L^p[0,\infty) \to L^p[0,\infty)$. When $p = \infty$, we get directly that $\|T_\infty\| = \|S_\infty\| = 1$. When $p = 1$ the operators are unbounded. This can be seen by considering $g_n = n\,1_{[0,1/n]}$. Then $\|g_n\|_1 = 1$, while via Tonelli (everything is non-negative)

$$\begin{aligned}\|T_1 g_n\|_1 &= n \int_0^1 \frac{1}{x} \int_0^x 1_{[0,1/n]}(t)\,dt\,dx = n \int_0^1 1_{[0,1/n]}(t) \int_t^1 \frac{1}{x}\,dx\,dt \\ &= -n \int_0^{1/n} \log t\,dt = 1 + \log n \xrightarrow[n\to\infty]{} \infty.\end{aligned}$$

That is why we do not consider the case $p = 1$.

For $p > 1$ we have, using the substitution $v = t/x$ and Minkowski's Integral Inequality (2.49),

$$\begin{aligned}\|T_p f\|_p &= \left( \int_0^1 \left| \frac{1}{x} \int_0^x f(t)\,dt \right|^p dx \right)^{1/p} = \left( \int_0^1 \left| \int_0^1 f(xt)\,dt \right|^p dx \right)^{1/p} \\ &\le \int_0^1 \left( \int_0^1 |f(xt)|^p\,dx \right)^{1/p} dt = \int_0^1 \left( \frac{1}{t} \int_0^t |f(x)|^p\,dx \right)^{1/p} dt \\ &\le \|f\|_p \int_0^1 t^{-1/p}\,dt = \frac{p}{p-1} \|f\|_p = q\,\|f\|_p,\end{aligned}$$

where as usual $q = p/(p-1)$ is the conjugate exponent to $p$ (with the natural interpretation as 1 when $p = \infty$). It is possible to show that

$$\|T_p\| = \|S_p\| = q.$$

It can be shown (Exercises 9.5.17 and 9.5.18) that

$$\sigma(T_p) = \left\{\lambda : \ \left|\lambda - \frac{q}{2}\right| \leq \frac{q}{2}\right\} = \overline{B_{q/2}(q/2)}$$

and

$$\sigma(S_p) = \left\{\lambda : \ \left|\lambda - \frac{q}{2}\right| = \frac{q}{2}\right\} = \partial B_{q/2}(q/2).$$

For $T_p$ and $p < \infty$, the point spectrum is precisely the whole interior of the disk, and each eigenvalue has multiplicity 1. For $T_\infty$, every nonzero element of the spectrum is an eigenvalue of multiplicity 1. In the case $p = 2$, it is not hard to show that $T_2^* T_2 = T_2 + T_2^* = T_2 T_2^*$, and it follows immediately that $T_2 = I - V$ for a unitary $V$.

### 9.5.1. Exercises.

(9.5.1) Let $T \in \mathcal{B}(\mathcal{X})$, $\alpha, \beta \in \mathbb{C}$. Show that $\sigma(\alpha T + \beta I) = \alpha\sigma(T) + \beta$.

(9.5.2) Let $\mathcal{X}$ be a Banach space with dimension (finite or infinite) at least 2. Given $a, b \in \mathcal{X}$ linearly independent, let $\mathcal{X}_0 = \operatorname{span}\{a, b\}$. Show that

- (i) there exists $\varphi \in \mathcal{X}^*$ with $\varphi(\alpha a + \beta b) = \beta$;
- (ii) there exists a bounded surjective projection $P : \mathcal{X} \to \mathcal{X}_0$;
- (iii) the linear operator $T : \mathcal{X} \to \mathcal{X}$ given by $Tx = \varphi(Px)a$ is bounded;
- (iv) $\sigma(T) = \{0\}$;
- (v) there exists $T \in \mathcal{B}(\mathcal{X})$ with $\|T\| = 1$ and $\operatorname{spr}(T) = 0$.

(9.5.3) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Show that if there exists nonzero $p \in \mathbb{C}[x]$ such that $p(T) = 0$ then $T$ admits an eigenvalue.

(9.5.4) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Suppose that $p \in \mathbb{C}[x]$ satisfies $p(T) = 0$. Show that there exists a root of $p$ that is an eigenvalue for $T$. Are all roots of $p$ eigenvalues of $T$? Provide roof/counterexample.

(9.5.5) Let $\mathcal{X}$ be an infinite-dimensional normed space. Show that there exists $T \in \mathcal{B}(\mathcal{X})$ such that $p(T) \neq 0$ for all $p \in \mathbb{C}[x]$.

(9.5.6) Let $\mathcal{X}$ be a normed space and $T \in \mathcal{B}(\mathcal{X})$. Suppose that $p(T) \neq 0$ for all $p \in \mathbb{C}[x]$. Does this imply that $T$ has no eigenvalues?

(9.5.7) Let $\mathcal{X}$ be a Banach space and $\mathcal{X}_1, \mathcal{X}_2 \subset \mathcal{X}$ subspaces with $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_1 \subset \mathcal{X}_1$ and $T\mathcal{X}_2 \subset \mathcal{X}_2$. Show that

$$\sigma(T) = \sigma(T|_{\mathcal{X}_1}) \cup \sigma(T|_{\mathcal{X}_2}).$$

(9.5.8) Write the details of Example 9.5.6.

(9.5.9) Show that $\sigma_p(T)$, $\sigma_r(T)$, and $\sigma_c(T)$ are mutually disjoint.

(9.5.10) Prove the relations (9.15), (9.16), and (9.17).

(9.5.11) Prove the relation (9.18).

(9.5.12) Construct an example where $\sigma_{ap}(T) \cap \sigma_r(T) \neq \varnothing$.

(9.5.13) In the context of Example 9.5.9 calculate explicitly $S(T-\lambda I)$ and $(T-\lambda I)S$ for $\lambda \notin [0,1]$.

(9.5.14) With $T$ the left unilateral shift as in Example 9.5.10, show that

$$\dim\ker(T-\lambda I) = 1$$

for all $\lambda \in \mathbb{D}$.

(9.5.15) Let $\mathcal{X} = L^2(-\infty,\infty)$ and let $T \in \mathcal{B}(\mathcal{X})$ the translation operator

$$(Tf)(x) = f(x+1).$$

Find the norm and the parts of the spectrum of $T$.

(9.5.16) Let $\mathcal{X} = L^2(0,\infty)$ and let $T \in \mathcal{B}(\mathcal{X})$ the translation operator

$$(Tf)(x) = f(x+1).$$

Find the norm and the parts of the spectrum of $T$.

(9.5.17) Let $p \in (1,\infty)$ and $T \in \mathcal{B}(L^p[0,1])$ be the Hardy operator as in Example 9.5.22. Show that $\|T\| = q$ and find $\sigma(T)$, $\sigma_p(T)$, $\sigma_r(T)$, $\sigma_c(T)$, and $\sigma_{ap}(T)$. *(This exercise is mostly computational, but nailing the right ideas and performing all the computations will possibly not be a trivial task; so this exercise and Exercise 9.5.18 should be seen more as a minor project rather than a couple of exercises)*

(9.5.18) Let $p \in (1,\infty)$ and $S \in \mathcal{B}(L^p[0,\infty))$ be the Hardy operator as in Example 9.5.22. Show that $\|S\| = q$ and find $\sigma(S)$, $\sigma_p(S)$, and $\sigma_{ap}(S)$. *(See the disclaimer in Exercise 9.5.17)*

(9.5.19) Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$, surjective. Let $\mathcal{Y} = \ker T$. Show that $\tilde{T} : \mathcal{X}/\mathcal{Y} \to \mathcal{X}$ given by $\tilde{T}(x+\mathcal{Y}) = Tx$ is linear, bijective, and bounded.

(9.5.20) Let $X$ be a compact Hausdorff space and $\psi : X \to X$ continuous. Let $T : C(X) \to C(X)$ be given by $Tf = f \circ \psi$. Show that

(i) $T$ is injective if and only if $\psi$ is surjective;

(ii) $T$ is surjective if and only if $\psi$ is injective.

(9.5.21) With the notation of Example 9.5.20, show that

$$f_\lambda(t) = \sum_{k=0}^{\infty} \lambda^k R^k g_0 = \sum_{k=0}^{\infty} \lambda^k \left(2^{k+1}t - \frac{1-2\lambda}{1-\lambda}\right) 1_{[2^{-k-1},2^{-k}]}(t).$$

## 9.6. Compact Operators

If we consider an arbitrary Banach space $\mathcal{X}$, a priori it looks hard to construct any linear operator other than the identity $I_{\mathcal{X}}$, given by $I_{\mathcal{X}} x = x$. But we can always do the following. By Hahn–Banach the dual $\mathcal{X}^*$ is non-trivial, so there exists nonzero $f \in \mathcal{X}^*$. Given nonzero $x_1 \in \mathcal{X}$ we can form the operator $Tx = f(x)x_1$; linearity is apparent, and it is bounded, for $\|Tx\| \leq \|f\| \, \|x_1\| \, \|x\|$. So $T \in \mathcal{B}(\mathcal{X})$. More generally, given $x_1, \ldots, x_n \in \mathcal{X}$ and $f_1, \ldots, f_n \in \mathcal{X}^*$, we can form the operator $T$ given by

$$Tx = \sum_{j=1}^{n} f_j(x) \, x_j. \tag{9.32}$$

This operator is bounded, since

$$\|Tx\| \leq \sum_{j=1}^{n} |f_j(x)| \, \|x_j\| \leq \left( \sum_{j=1}^{n} \|f_j\| \, \|x_j\| \right) \|x\|.$$

We can even repeat the trick with infinitely many points and operators, as long as $\sum_j \|f_j\| \, \|x_j\| < \infty$ to guarantee convergence. In the case of (9.32), since $\operatorname{ran} T \subset \operatorname{span}\{x_1, \ldots, x_n\}$, it is finite-dimensional. We call such $T$ a **finite-rank operator**. When $\mathcal{X}$ is infinite-dimensional operators which are not of finite-rank always exist, starting with the identity $I_{\mathcal{X}}$.

**9.6.1. Definition.**

*A linear operator $T : \mathcal{X} \to \mathcal{Y}$ is* ***compact*** *if $\overline{T(B_1^{\mathcal{X}}(0))}$ is compact in $\mathcal{Y}$. We will denote by $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ the set of compact operators $\mathcal{X} \to \mathcal{Y}$, and we will use $\mathcal{K}(\mathcal{X})$ when $\mathcal{Y} = \mathcal{X}$.*

*When $T$ is bounded and $\dim \operatorname{ran} T < \infty$, we say that $T$ is* ***finite-rank****. We denote the set of finite-rank operators $\mathcal{X} \to \mathcal{Y}$ by $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ and we use $\mathcal{F}(\mathcal{X})$ when $\mathcal{Y} = \mathcal{X}$.*

We need to require that $T$ is bounded in the definition of finite-rank, because unbounded operators of finite rank do exist. We saw already on page 331 that unbounded linear functionals exist, and these give examples of unbounded operators of rank one.

Here is some basic information about compact operators.

***9.6.2. Proposition.***

*A compact operator is bounded. Finite-rank operators are compact. Limits of compact operators are compact. When $\mathcal{X} = \mathcal{Y}$, the identity operator is compact if and only if* $\dim \mathcal{X} < \infty$.

*Proof.* A compact set in a metric space is bounded. So if $T$ is compact there exists $c > 0$ such that $\|Tx\| \le c$ for all $x \in B_1^{\mathcal{X}}(0)$. That is, $\|Tx\| \le c\|x\|$ for all $x \in \mathcal{X}$.

When $T$ is finite-rank, $\overline{T(B_1^{\mathcal{X}}(0))}$ is contained in a ball in a finite-dimensional space, and so it is compact by Theorem 5.2.9.

Now suppose that $T = \lim_n T_n$, with $T_n$ compact for all $n$. Fix $\varepsilon > 0$; then there exists $n$ such that $\|T - T_n\| < \varepsilon/3$. As $T_n$ is compact, the image of the unit ball admits an $\varepsilon/2$-net; that is, there exist $y_1, \dots, y_m \in \mathcal{Y}$ such that $\overline{T_n B_1^{\mathcal{X}}(0)} \subset \bigcup_{k=1}^m B_{\varepsilon/2}(y_k)$. For any $x \in B_1^{\mathcal{X}}(0)$,

$$\|Tx - y_k\| \le \|Tx - T_n x\| + \|T_n x - y_k\| \le \frac{\varepsilon}{3} + \frac{\varepsilon}{2} < \varepsilon.$$

So $\overline{TB_1^{\mathcal{X}}(0)}$ can be covered by finitely many balls of radius $\varepsilon$. As $\varepsilon$ was arbitrary, we have shown that $\overline{TB_1^{\mathcal{X}}(0)}$ admits an $\varepsilon$-net and is thus compact. So $T$ is compact.

If $I_{\mathcal{X}}$ is compact, then every sequence in $\overline{B_1^{\mathcal{X}}(0)}$ admits a convergent subsequence. So $\overline{B_1^{\mathcal{X}}(0)}$ is compact, and thus $\dim \mathcal{X} < \infty$ by Theorem 5.2.9. Conversely, when $\dim \mathcal{X} < \infty$ the closed unit ball is compact, and so $I_{\mathcal{X}}$ is compact. $\square$

Both $\mathcal{F}(\mathcal{X})$ and $\mathcal{K}(\mathcal{X})$ are bilateral ideals in $\mathcal{B}(\mathcal{X})$. More generally, if $S \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ and $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$, then $TS \in \mathcal{K}(\mathcal{Y})$ and $ST \in \mathcal{K}(\mathcal{X})$; similarly, if $T \in \mathcal{F}(\mathcal{X}, \mathcal{Y})$, then $TS \in \mathcal{F}(\mathcal{Y})$ and $ST \in \mathcal{F}(\mathcal{X})$ (Exercises 9.6.5 and 9.6.6).

We showed in Proposition 9.6.2 that $\overline{\mathcal{F}(\mathcal{X}, \mathcal{Y})} \subset \mathcal{K}(\mathcal{X}, \mathcal{Y})$. When $\mathcal{X}$ is a Hilbert space—and, more generally, when $\mathcal{X}$ has a Schauder basis (or, even more generally, it has the approximation property)—this inclusion is an equality (Proposition 9.8.5).

***9.6.3. Lemma.***

*Let $T \in \mathcal{K}(\mathcal{X})$ with $\mathcal{X}$ infinite-dimensional. Then $T$ is not surjective.*

*Proof.* If $T$ is surjective, by the Open Mapping Theorem (6.3.5) the map $T$ is open. Then the compact set $\overline{T(B_1^{\mathcal{X}}(0))}$ contains a ball $B_r(0)^{\mathcal{X}}$ for some $r > 0$. Thus $\overline{B_r^{\mathcal{X}}(0)}$ is compact, and by scaling $\overline{B_1^{\mathcal{X}}(0)}$ is compact. Then $\dim \mathcal{X} < \infty$

by Theorem 5.2.9. As we were assuming that $\mathcal{X}$ is infinite-dimensional, $T$ cannot be surjective. □

One usually treats compact operators not by using the above definition but some equivalent, more tractable property. The properties we consider here all agree for compact operators on Hilbert spaces (Proposition 10.6.4).

**9.6.4. Definition.**

*An operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is* ***completely continuous*** *if for every sequence $\{x_n\} \subset \mathcal{X}$ such that $x_n \xrightarrow{\text{weak}} x$ (weak convergence), we have that $Tx_n \to Tx$ (in norm).*

Note the requirement "sequence" in the definition. As we see below, the notion of complete continuity agrees with compactness when $\mathcal{X}$ is reflexive (for instance, when $\mathcal{X}$ is a Hilbert space, or when $\mathcal{X} = \ell^p(\mathbb{N})$, $1 < p < \infty$), but not in general.

***9.6.5. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$.*

*(i) If $T$ is compact, then it is completely continuous. If $\mathcal{X}$ is reflexive and $T$ is completely continuous, then $T$ is compact.*

*(ii) If $T^*$ is compact (this is equivalent to $T$ compact, as established in Theorem 9.6.12), then $T^*$ maps weak$^*$-convergent sequences to norm-convergent sequences.*

*Proof.* Suppose that $T$ is compact and that $x_n \xrightarrow{\text{weak}} 0$. By Proposition 7.1.11, there exists $c > 0$ with $\|x_n\| < c$ for all $n$. Then $\{Tx_n\} \subset \overline{TB_c(0)}$, which is compact. So there exists $\{x_{n_k}\}$ and $y \in Y$ such that $Tx_{n_k} \to y$. As $x_{n_k} \xrightarrow{\text{weak}} 0$, for any $f \in \mathcal{Y}^*$ we have $f(y) = \lim_k f(Tx_{n_k}) = \lim_k (T^*f)(x_{n_k}) = 0$; thus $y = 0$. From $x_n \xrightarrow{\text{weak}} 0$, we have $f(Tx_n) = (T^*f)(x_n) \to 0$, so $Tx_n \xrightarrow{\text{weak}} 0$. So the only possible accumulation point of $\{Tx_n\}$ is 0. This forces $Tx_n \to 0$ (otherwise, take a subsequence that is far from zero, and by compactness it has a convergent subsequence). So $T$ is completely continuous.

The proof when $T^*$ is compact runs in an analogous way, using Proposition 7.2.6 instead (Proposition 7.1.11).

Now assume that $\mathcal{X}^{**} = \mathcal{X}$, and that $T$ is completely continuous. Suppose first that $\mathcal{X}$ is separable. By Proposition 7.1.15 $(\overline{B_1^{\mathcal{X}}(0)}, \sigma(\mathcal{X}, \mathcal{X}^*))$ is a separable compact metric space. Let $\{Tx_n\} \subset TB_1^{\mathcal{X}}(0)$, with $\|x_n\| \leq 1$. Since $\overline{B_1^{\mathcal{X}}(0)}$ is compact and metric, there exists a subsequence $\{x_{n_k}\}$ and

$x \in \mathcal{X}$ with $x_{n_k} \xrightarrow{\text{weak}} x$. We are assuming that $T$ is completely continuous, so $Tx_{n_k} \to Tx$. Thus $TB_1^{\mathcal{X}}(0)$ is (sequentially) pre-compact and hence pre-compact (see Remark 1.8.20), and so $T$ is compact.

If we now take a general $\mathcal{X}$, given $\{x_n\} \subset B_1^{\mathcal{X}}(0)$, we may take $\mathcal{X}_0 = \overline{\operatorname{span}\{x_n\}}$ which is Banach, separable, and reflexive (Proposition 7.2.16). So $T_1 = T|_{\mathcal{X}_0}$ is compact, and $\{Tx_n\} = \{T_1x_n\}$ has a convergent sequence; implying that $T$ is compact. □

*9.6.6. Remark*. For an example of a completely continuous operator that is not compact, let $\mathcal{X} = \ell^1(\mathbb{N})$ and let $T \in \mathcal{B}(\mathcal{X})$ be any non-compact operator (we can take $T = I_{\mathcal{X}}$ if no other operator is available). Suppose that $x_n \xrightarrow{\text{weak}} 0$. By Proposition 7.1.22, $x_n \to 0$; as $T$ is bounded, $Tx_n \to 0$. So $T$ is completely continuous. More generally, if $\mathcal{X}, \mathcal{Y}$ are Banach spaces with at least one of them equal to $\ell^1(\mathbb{N})$, then every $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is completely continuous (Exercise 9.6.1).

*9.6.7. Remark*. There exist Banach spaces which are not reflexive and where every completely continuous operator is compact. Concretely, $\mathcal{X} = c_0$ is one such example. Here is a sketch of an argument. There are non-completely continuous operators in $\mathcal{B}(c_0)$. Namely, the canonical basis satisfies $e_n \xrightarrow{\text{weak}} 0$ but it does not converge in norm; hence the identity operator is not completely continuous. The set of completely continuous operators is a normed closed ideal of $\mathcal{B}(\mathcal{X})$. And $\mathcal{K}(c_0)$ is also a closed ideal. Finally, one shows that $\mathcal{B}(c_0)$ admits a unique non-trivial ideal (Exercise 9.8.16).

Here is an interesting consequence of Proposition 9.6.5:

**9.6.8. Corollary.**

*Let $\mathcal{H}$ be a Hilbert space and $T : \mathcal{H} \to \ell^1(\mathbb{N})$ be linear and bounded. Then $T$ is compact.*

*Proof.* Suppose that $\xi_n \xrightarrow{\text{weak}} \xi$ in $\mathcal{H}$. For any $f \in \ell^1(\mathbb{N})^*$, since $T$ is bounded we have that $f \circ T \in \mathcal{H}^*$. Then $f(T\xi_n)) \to f(T\xi)$ by definition of weak convergence. As this can be done for any $f \in \ell^1(\mathbb{N})^*$, this implies that $T\xi_n \xrightarrow{\text{weak}} T\xi$. By the fact that $\ell^1(\mathbb{N})$ has the Schur property (Proposition 7.1.22), we get that $T\xi_n \to T\xi$. Then $T$ is compact by Proposition 9.6.5 since $\mathcal{H}$ is reflexive. □

*9.6.9. Remark*. Corollary 9.6.8 is a particular case of Pitt's Theorem (9.8.16)

*9.6.10. Remark* . Corollary 9.6.8 allows us to produce another example of a surjective linear operator on a Banach space without right-inverse (we had a first example in Remark 6.2.5). By Theorem 6.1.10 there exists a surjective bounded linear map $T:\ell^1(\mathbb{N})\to\ell^2(\mathbb{N})$ (namely, $\ell^2(\mathbb{N})$ is isometrically isomorphic to a quotient of $\ell^1(\mathbb{N})$; $T$ is the composition of this isomorphism with the quotient map). Suppose that $S$ is a right-inverse for $T$, that is $S:\ell^2(\mathbb{N})\to\ell^1(\mathbb{N})$ bounded with $TS=\mathrm{id}_{\ell^2(\mathbb{N})}$. By Corollary 9.6.8 we have that $S$ is compact. By Exercise 9.6.6, $\mathrm{id}_{\ell^2(\mathbb{N})}=TS$ is compact, a contradiction since this would imply that $\ell^2(\mathbb{N})$ is finite-dimensional (Proposition 9.6.2).

We will include another result in the same vein. It is a particular case of Pitt's Theorem; but it is needed for its proof.

***9.6.11. Proposition.***

*Let $T\in\mathcal{B}(c_0,\ell^1(\mathbb{N}))$. Then $T$ is compact.*

*Proof.* Let $\{x_n\}\subset B_1^{c_0}(0)$. We want to show that $\{Tx_n\}\subset\ell^1(\mathbb{N})$ admits a convergent subsequence. Since $\ell^1(\mathbb{N})=c_0^*$ is separable, the weak topology is metrizable in the unit ball (Proposition 7.1.15). Thinking of $x_n$ as $\hat x_n$ in the double-dual $\ell^\infty(\mathbb{N})$, we have by Banach–Alaoglu (Theorem 7.2.13) that $\{\hat x_n\}$ admits a convergent subnet. Since $\hat x_n(f)=f(x_n)$, the weak$^*$-topology of $B_1^{c_0}(0)$ as a subset of $B_1^{\ell^\infty(\mathbb{N})}(0)$ agrees with the weak topology. As the latter is metrizable, there exists a weakly Cauchy subsequence $\{x_{n_j}\}$. A basic weak neighbourhood $N$ of $0$ in $\ell^1(\mathbb{N})$ is given by some and $g_1,\ldots,g_m\in\ell^\infty(\mathbb{N})$ so that

$$N=\{y\in\ell^1(\mathbb{N}):\ |g_s(y)|<1,\ s=1,\ldots,m\}.$$

As $T$ is bounded, $g_s\circ T\in c_0^*$. Let $M$ be the weak-neighbourhood of $0$ in $c_0$ given by

$$M=\{x\in c_0:\ |g_s(Tx)|<1,\ s=1,\ldots,m\}.$$

As $\{x_{n_j}\}$ is weak-Cauchy, there exists $j_0$ such that $x_{n_j}-x_{n_k}\in M$ for all $j,k\geq j_0$. That is, $|g_s(Tx_{n_j}-Tx_{n_k})|<\varepsilon$, which means that $Tx_{n_j}-Tx_{n_k}\in N$. This shows that $\{Tx_{n_j}\}$ is weak-Cauchy in $\ell^1(\mathbb{N})$. By the Schur property (Proposition 7.1.22) the sequence $\{Tx_{n_j}\}$ is norm-convergent. ☐

**9.6.12. Theorem. (Schauder)**

*Let $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$. Then $T$ is compact if and only $T^*$ is compact.*

*Proof.* Suppose that $T$ is compact. To show that $T^*$ is compact, take $\{g_n\} \subset B_1^{\mathcal{Y}^*}(0)$. By Banach–Alaoglu (Theorem 7.2.13), there exists $\{g_{n_k}\}$ and $g \in \mathcal{Y}^*$ such that $g_{n_k} \xrightarrow{\text{weak}^*} g$. Fix $\varepsilon > 0$. As $\overline{TB_1^{\mathcal{X}}(0)}$ is compact, there exist $y_1,\ldots,y_m \in \mathcal{Y}$ such that $TB_1^{\mathcal{X}}(0) \subset \bigcup_{j=1}^n B_{\varepsilon/3}(y_j)$. Given any $N \geq 1$, there exists $k$ such that $n_k \geq N$ and $|g_{n_k}(y_j) - g(y_j)| < \varepsilon/3$, $j = 1,\ldots,n$. If $x \in B_1^{\mathcal{X}}(0)$, there exists $j$ such that $\|Tx - y_j\| < \varepsilon/3$. Then

$$\begin{aligned}|(T^*g_{n_k} - T^*g)(x)| &= |(g_{n_k} - g)(Tx)| \leq |(g_{n_k} - g)(Tx - y_j)| + (g_{n_k} - g)(y_j)| \\ &\leq 2\|Tx - y_j\| + \frac{\varepsilon}{3} \leq \frac{3\varepsilon}{3} = \varepsilon.\end{aligned}$$

As this works for any $x \in B_1^{\mathcal{X}}(0)$, we have shown that $\|T^*g_{n_k} - T^*g\| < \varepsilon$, and so $T^*g_{n_k} \to T^*g$. Thus $\{T^*g_n\}$ has a convergent subsequence, and so $T^*$ is compact.

Now suppose that $T^*$ is compact. By the above, $T^{**}$ is compact. Then $T = T^{**}|_{\mathcal{X}}$ is compact. Indeed, if $\{x_n\} \subset B_1^{\mathcal{X}}(0)$ then $\{T^{**}\hat{x}_n\}$ admits a convergent subsequence. That makes $\{T^{**}\hat{x}_{n_k}\}$ Cauchy and so $\{Tx_{n_k}\}$ is Cauchy by Exercise 9.4.5. So $T$ is compact. □

When $\dim \mathcal{X} = \infty$, the algebra $\mathcal{K}(\mathcal{X})$ is not unital. So in principle we could consider the spectrum of a compact operator $T$ both in $\mathcal{B}(\mathcal{X})$ and in $\mathcal{A} = \mathcal{K}(\mathcal{X}) + \mathbb{C}I$. It turns out that though in general the spectrum does depend on the algebra, it does not in this situation. We know that since $\mathcal{A} \subset \mathcal{B}(\mathcal{X})$, we have $\sigma(T) \subset \sigma_{\mathcal{A}}(T)$ and $\partial\sigma_{\mathcal{A}}(T) \subset \partial\sigma(T)$ (Proposition 9.2.8). We will show in Theorem 9.6.13 that for $T$ compact $\sigma(T)$ is always countable, with at most one accumulation point. Then $\partial\sigma(T) = \sigma(T)$. If $\sigma_{\mathcal{A}}(T)$ had interior, then $\partial\sigma_{\mathcal{A}}(T)$ would be uncountable (take a ball in the interior, and then all rays from the centre have to touch the boundary at different points), contradicting $\partial\sigma_{\mathcal{A}}(T) \subset \partial\sigma(T)$. Thus

$$\sigma_{\mathcal{A}}(T) = \partial\sigma_{\mathcal{A}}(T) \subset \partial\sigma(T) = \sigma(T).$$

and then $\sigma_{\mathcal{A}}(T) = \sigma(T)$. In particular, we have proven that if $T \in \mathcal{K}(\mathcal{X})$ and $T + \lambda I$ is invertible, then $(T + \lambda I)^{-1} = S + \lambda^{-1}I$ for some $S \in \mathcal{K}(\mathcal{X})$.

**9.6.13. Theorem.**

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{K}(\mathcal{X})$. Then exactly one of the following holds:*

*(i)* $\sigma(T) = \{0\}$;

(ii) $\sigma(T) = \{0, \lambda_1, \dots, \lambda_n\}$, *where each nonzero* $\lambda_j$ *is an eigenvalue for* $T$, *and* $\dim\ker(T - \lambda_j I) < \infty$*;*

(iii) $\sigma(T) = \{0\} \cup \{\lambda_j\}$, *where each nonzero* $\lambda_j$ *is an eigenvalue for* $T$, $\lambda_j \to 0$, *and* $\dim\ker(T - \lambda_j I) < \infty$.

*In particular,* $\sigma(T) = \{0\} \cup \sigma_p(T)$.

*Proof.* We proceed by stating and proving several claims, that together imply the theorem.

**Claim 1.** If $\lambda \neq 0$ and $\ker(T - \lambda I) = \{0\}$, then $\operatorname{ran}(T - \lambda I)$ is closed.

Indeed, suppose that $T - \lambda I$ is not bounded below. Then there exist $\{x_n\}$ with $\|x_n\| = 1$ and $Tx_n - \lambda x_n \to 0$. As $T$ is compact, there exist $\{x_{n_k}\}$ and $y$ such that $Tx_{n_k} \to y$. Since

$$\|y\| = \lim_k \|Tx_{n_k}\| = \lim_k \|(T - \lambda I)x_{n_k} + \lambda x_{n_k}\| \to \lambda,$$

we get that $y \neq 0$. Knowing that $Tx_{n_k}$ converges to $y$, we get that $\lambda x_{n_k}$ also converges to $y$; then

$$Ty = \lim_k \lambda T x_{n_k} = \lambda y$$

and so $y \in \ker(T - \lambda I)$, a contradiction. Thus $T - \lambda I$ is bounded below; by Proposition 9.5.4 we have that $\operatorname{ran}(T - \lambda I)$ is closed.

**Claim 2.** If $\lambda \neq 0$ and $\lambda \in \sigma(T)$, then $\lambda \in \sigma_p(T)$ or $\lambda \in \sigma_p(T^*)$.

Suppose that $\lambda \notin \sigma_p(T) \cup \sigma_p(T^*)$. From $\lambda \notin \sigma_p(T)$ we get that $\ker(T - \lambda I) = \{0\}$; by Claim 1, $\operatorname{ran}(T - \lambda I)$ is closed. By Proposition 9.5.4, $T - \lambda I$ is bounded below. As $\ker(T^* - \lambda I) = \{0\}$, we get from (9.13) in Proposition 9.4.6 that $\overline{\operatorname{ran}(T - \lambda I)} = \mathcal{X}$; but $\operatorname{ran}(T - \lambda I)$ is closed, so $T - \lambda I$ is surjective. Then $T - \lambda I$ is bijective, and by the Inverse Mapping Theorem 6.3.6 it is invertible, a contradiction.

**Claim 3.** If $\{\lambda_n\} \subset \sigma_p(T)$ are distinct, then $\lambda_n \to 0$.

For each $n \in \mathbb{N}$, choose $x_n \in \ker(T - \lambda_n I)$ nonzero. Let

$$M_n = \operatorname{span}\{x_1, \dots, x_n\}.$$

If $\sum_{j=1}^n c_j x_j = 0$, then $0 = \sum_j c_j T x_j = \sum_j c_j \lambda_j x_j$. And we may iterate this, so $\sum_j c_j \lambda_j^k x_j = 0$ for all $k$. Taking linear combinations, $\sum_{j=1}^n c_j p(\lambda_j) x_j = 0$ for all polynomials $p$. Choosing $p$ such that $p(\lambda_j) = 1$ and $p(\lambda_k) = 0$ for $k \neq j$, we get that $c_1 = \cdots = c_n = 0$. Then $x_1, \dots, x_n$ are linearly independent, and $\dim M_n = n$. This shows that the inclusion $M_n \subset M_{n+1}$ is proper. By Riesz's Lemma (5.2.8) there exists $y_n \in M_n$ with $\|y_n\| = 1$ and $\operatorname{dist}(y_n, M_{n-1}) > 1/2$. Write $y_n = \sum_{j=1}^n \alpha_j x_j$. Then

$$(T - \lambda_n I)y_n = \sum_{j=1}^n \alpha_j (T - \lambda_n I)x_j = \sum_{j=1}^{n-1} \alpha_j (T - \lambda_n I)x_j$$

$$= \sum_{j=1}^{n-1} \alpha_j(\lambda_j - \lambda_n)x_j \in M_{n-1}.$$

If $n > m$,

$$T\left(\frac{1}{\lambda_n}y_n\right) - T\left(\frac{1}{\lambda_m}y_m\right) = \overbrace{\frac{1}{\lambda_n}(T-\lambda_n I)y_n}^{\in M_{n-1}} - \overbrace{\frac{1}{\lambda_m}(T-\lambda_m I)y_m}^{\in M_m \subset M_{n-1}} + y_n - \overbrace{y_m}^{\in M_m \subset M_{n-1}}.$$

So

$$\left\| T\left(\frac{1}{\lambda_n}y_n\right) - T\left(\frac{1}{\lambda_m}y_m\right)\right\| \geq \operatorname{dist}(y_n, M_{n-1}) > \frac{1}{2}.$$

Therefore $\{T(\frac{1}{\lambda_n}y_n)\}$ admits no convergent subsequence. As $T$ is compact, it follows that $\{\frac{1}{\lambda_n}y_n\}$ admits no bounded subsequence. Thus

$$\frac{1}{|\lambda_n|} = \left\|\frac{1}{\lambda_n}y_n\right\| \to \infty,$$

which implies that $\lambda_n \to 0$.

**Claim 4.** If $\lambda \in \sigma(T)\setminus\{0\}$, then $\lambda$ is isolated. In particular, $\sigma(T)$ is countable by Exercise 1.8.8.

If $\lambda_n \to \lambda$, with $\lambda_n \in \sigma(T)$ for all $n$, eventually $\lambda_n \neq 0$. By Claim 2, $\lambda_n \in \sigma_p(T) \cup \sigma_p(T^*)$. So there exists $\{\lambda_{n_k}\}$ such that it is contained either in $\lambda_p(T)$ or $\lambda_p(T^*)$. By Claim 3 (since $T^*$ is compact by Theorem 9.6.12), $\lambda_{n_k} \to 0$, a contradiction.

**Claim 5.** If $\lambda \in \sigma_p(T)$, $\lambda \neq 0$, then $\dim\ker(T-\lambda I) < \infty$.

Let $M_\lambda = \ker(T-\lambda I)$, and let $T_\lambda = T|_{M_\lambda}$. Then $T_\lambda x = \lambda x$ for all $x \in M_\lambda$, so $T_\lambda \in \mathcal{B}(M_\lambda)$. That is, $T_\lambda = \lambda I$. Thus $\lambda B_1^{M_\lambda}(0) = T_\lambda B_1^{M_\lambda}(0) \subset TB_1^{\mathcal{X}}(0)$, pre-compact. Then $\dim M_\lambda < \infty$ by Theorem 5.2.9.

**Claim 6.** If $\lambda \in \sigma(T) \setminus \{0\}$, then $\lambda \in \sigma_p(T)$.

Since $\lambda$ is isolated (Claim 4), by Proposition 9.3.4 (applied to $T$ on the Banach algebra $\mathcal{B}(\mathcal{X})$) we get that $\lambda$ is an eigenvalue and we get the nonzero projection $E_\lambda$ onto the eigenspace $M_\lambda$ of $\lambda$. From $TE_\lambda = \lambda E_\lambda$ we have that $T|_{M_\lambda} = \lambda I_{M_\lambda}$ and $T|_{M_\lambda} \in \mathcal{B}(M_\lambda)$. So $T|_{M_\lambda}$ is invertible; and being the restriction of a compact operator, it is also compact. This forces $\dim M_\lambda < \infty$ by Theorem 5.2.9. □

***9.6.14. Corollary. (The Fredholm Alternative)***

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{K}(\mathcal{X})$, $\lambda \in \mathbb{C} \setminus \{0\}$. Then either $0 < \dim\ker(T-\lambda I) < \infty$, or $T - \lambda I$ is invertible.*

*Proof.* In all possibilities given by Theorem 9.6.13, a nonzero element of the spectrum is an eigenvalue with finite-dimensional eigenspace. □

Historically, the Fredholm Alternative appeared in the context of solving certain differential equations. If we have an equation of the form $Tx = \lambda x + b$, with $\lambda \in \mathbb{C} \setminus \{0\}$ and $T \in \mathcal{K}(\mathcal{X})$, then Corollary 9.6.14 says that the equation behaves like linear systems in linear algebra, which means that one of the following occurs:

- the system has a unique solution (when $\lambda \notin \sigma(T)$); or
- if the system is compatible, it has $\dim\ker(T - \lambda I)$ solutions.

*9.6.15. Example.* Theorem 9.6.13 shows that compact operators can be expected to be somehow a generalization of matrices. The analogy is not perfect, though. Let $\mathcal{X} = L^2[0,1]$ and consider the **Volterra operator** $V \in \mathcal{B}(\mathcal{X})$ given by

$$(Vf)(s) = \int_0^s f(t)\,dt.$$

It is not hard to show that $V$ is compact (see Exercise 9.6.2). In particular any nonzero $\lambda \in \sigma(V)$ would have to be an eigenvalue by Theorem 9.6.13. Now, if $\lambda \neq 0$ and $Vf = \lambda f$, we get

$$f(x) = \frac{1}{\lambda} \int_0^x f(t)\,dt. \tag{9.33}$$

Using that $f$ is in $L^2$ we have, for $x < y$,

$$\begin{aligned} |f(y) - f(x)| &= \frac{1}{\lambda} \left| \int_x^y f(t) dt \right| \leq \frac{1}{\lambda} \int_x^y |f(t)| dt = \frac{1}{\lambda} \int_0^1 |f(t)|\, 1_{[x,y]}(t)\,dt \\ &\leq \frac{\|f\|_2}{\lambda} \left( \int_0^1 (1_{[x,y]})^2\,dt \right)^{1/2} = \frac{\|f\|_2}{\lambda}\sqrt{y - x}. \end{aligned}$$

So $f$ is continuous. Then looking at (9.33) again we now know that $f$ is differentiable; and after differentiation, (9.33) is the equation $f = \lambda f'$. This implies that $f(t) = c\,e^{t/\lambda}$. But $f(0) = 0$ (again from (9.33)), so the only solution for (9.33) is $f = 0$, and $\lambda$ cannot be an eigenvalue.

When $\lambda = 0$, if $Vf = 0$, then $f = 0$. So $V$ has no eigenvalues, and $\sigma(V) = \{0\}$. The Volterra operator is an example of a **quasinilpotent operator**. See Exercise 9.6.4 for another simple example of a quasinilpotent operator.

We will see in Section 10.6 that for good enough—that is, normal—compact operators acting on a Hilbert space, the analogy with matrices works a lot better.

### 9.6.1. Exercises.

(9.6.1) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces with at least one of them equal to $\ell^1(\mathbb{N})$, and let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that $T$ is completely continuous.

(9.6.2) Let $(X, \mathcal{A}, \mu)$ be a measure space with finite measure, $1 < p < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. If $k : X \times X \to \mathbb{C}$ is an $\mathcal{A} \boxdot \mathcal{A}$-measurable function such that

$$\sup \left\{ \int_X |k(x,y)|^q \, d\mu(y) \, : \, x \in X \right\} < \infty,$$

show that

$$(Kf)(x) = \int_X k(x,y) f(y) \, d\mu(y)$$

defines a compact operator on $L^p(\mu)$. *(Hint: use complete continuity and reflexivity)*

(9.6.3) Consider the multiplication operator $M_b$ as in Example 9.5.6. Show that $M_b$ is compact if and only if $b \in c_0$.

(9.6.4) Let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the linear operator induced by

$$Te_n = \tfrac{1}{n+1} \, e_{n+1}, \qquad n \in \mathbb{N}.$$

(i) Show that $T$ is bounded on $c_{00}$, so that it extends to all of $\ell^2(\mathbb{N})$ and is bounded.

(ii) Show that $T$ is compact.

(iii) Show that $T$ is quasinilpotent.

(9.6.5) Show that $\mathcal{F}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{K}(\mathcal{X}, \mathcal{Y})$ are subspaces of $\mathcal{B}(\mathcal{X}, \mathcal{Y})$.

(9.6.6) Show that if $R \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$ and $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$, then $TR \in \mathcal{K}(\mathcal{Z}, \mathcal{Y})$. And if $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ then $ST \in \mathcal{K}(\mathcal{X}, \mathcal{Z})$. Show also that analog results hold with $T \in \mathcal{F}(\mathcal{X}, \mathcal{Y})$.

(9.6.7) Let $\mathcal{X}$ be a normed space and $\mathcal{J}$ a (not necessarily closed) nonzero ideal. Show that $\mathcal{F}(\mathcal{X}) \subset \mathcal{J}$.

(9.6.8) Let $T \in \mathcal{B}(\mathcal{X})$. Show that if $\operatorname{ran} T$ is finite-dimensional, then $T$ is of the form (9.32).

(9.6.9) Let $\mathcal{X}, \mathcal{Y}$ be infinite-dimensional Banach spaces, $T \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. Show that $T$ is not bounded below on any infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$.

(9.6.10) Let $V \in \mathcal{B}(L^2[0,1])$ be the Volterra operator defined in Example 9.6.15,

$$(Vf)(x) = \int_0^x f.$$

As mentioned, knowing that $V$ is compact and has no eigenvalues, it follows immediately that $\sigma(V) = \{0\}$. Prove this fact explicitly by calculating $(T - \lambda I)^{-1}$ for any $\lambda \neq 0$.

(9.6.11) Let $\mathcal{X} = \ell^2(\mathbb{N})$ and $\{\lambda_n\} \subset \mathbb{C}$ with $\lambda_n \to 0$. Show that there exists $T \in \mathcal{K}(\mathcal{X})$ with $\sigma(T) = \{\lambda_n\}$.

(9.6.12) Let $\mathcal{X}$ be a normed space and $E \in \mathcal{B}(\mathcal{X})$ an idempotent. Show that $E$ has finite-rank if and only if it is compact.

(9.6.13) Let $\mathcal{X}$ be a normed space and $\mathcal{X}_0 \subset \mathcal{X}$ a nonzero finite-dimensional proper subspace. Show that there are uncountably many idempotents $E \in \mathcal{B}(\mathcal{X})$ with $X_0 = E\mathcal{X}$.

(9.6.14) Let $\mathcal{X}$ be a Banach space and $E \in \mathcal{B}(\mathcal{X})$ an idempotent. Show that $\operatorname{ran} E$ and $\operatorname{ran}(I - E)$ are closed.

(9.6.15) Let $E \in \mathcal{B}(\mathcal{X})$ be a finite-rank idempotent. Show that $E^*$ is a finite-rank idempotent and that $\dim \operatorname{ran} E = \dim \operatorname{ran} E^*$.

(9.6.16) Let $\mathcal{X}$ be a vector space and $P, Q : \mathcal{X} \to \mathcal{X}$ be idempotents. Show that if $\ker Q \subset \ker P$ and $\operatorname{ran} Q \subset \operatorname{ran} P$, then $P = Q$.

## 9.7. Fredholm Operators

The study of the spectrum of $T \in \mathcal{K}(\mathcal{X})$ entailed looking at operators of the form $T - \lambda I$ for $T$ compact. It turns out that such operators are examples of a very important class of operators, the Fredholm Operators. Before defining them we will establish a couple of useful results.

***9.7.1. Lemma.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ with closed range. Then*

$$\dim \ker T^* = \dim \big(\mathcal{Y}/\operatorname{ran} T\big)^*. \tag{9.34}$$

*Proof.* By (9.11) in Proposition 9.4.6 and Theorem 7.3.7,

$$\ker T^* = (\operatorname{ran} T)^o \simeq \Big(\mathcal{Y}/\operatorname{ran} T\Big)^*,$$

so their dimensions are equal. □

***9.7.2. Proposition.***

*Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{K}(\mathcal{X})$. Then*

*(i)* $\dim \ker(I - T) < \infty$;

*(ii)* $\operatorname{ran}(I-T)$ *is closed;*

*(iii)* $\operatorname{ran}(I-T) = \ker(I-T^*)_o$*;*

*(iv)* $\ker(I-T) = \{0\}$ *if and only if* $\operatorname{ran}(I-T) = \mathcal{X}$*;*

*(v)* $\dim \mathcal{X}/\operatorname{ran}(I-T) = \dim\ker(I-T^*)$*;*

*(vi)* $\dim\ker(I-T) = \dim\ker(I-T^*)$.

*Proof.* (i) Take $\lambda = 1$ in Corollary 9.6.14.

(ii) Let $\{y_k\} \subset \operatorname{ran}(I-T)$ be Cauchy; write $y = \lim_k y_k$. We have $y_k = x_k - Tx_k$ for some $\{x_k\} \subset \mathcal{X}$. Since $T$ is compact, the sequence $\{Tx_k\}$ admits a convergent subsequence $\{Tx_{k_j}\}$. Let $y_0 = \lim Tx_{k_j}$. Now

$$x_{k_j} = y_{k_j} + Tx_{k_j} \xrightarrow[j\to\infty]{} y + y_0$$

is convergent. Then, if we write $x_0 = \lim_j x_{k_j}$,

$$y = \lim_j x_{k_j} - y_0 = \lim_j x_{k_j} - Tx_{k_j} = x_0 - Tx_0 \in \operatorname{ran}(I-T),$$

so $\operatorname{ran}(I-T)$ is closed.

(iii) Knowing that $\operatorname{ran}(I-T)$ is closed, we get $\operatorname{ran}(I-T) = \ker(I-T^*)_o$ by (9.13) in Proposition 9.4.6.

(iv) By Theorem 9.6.12, $T^* \in \mathcal{K}(\mathcal{X}^*)$. By Proposition 9.5.1, $\sigma(T^*) = \sigma(T)$. If $\ker(I-T) = \{0\}$, by Theorem 9.6.13 we have that $1 \notin \sigma(T)$. So $1 \notin \sigma(T^*)$, and $I - T^*$ is invertible. In particular, $\ker(I-T^*) = \{0\}$, and then $\operatorname{ran}(I-T) = \mathcal{X}$ by (iii). Conversely, if $\operatorname{ran}(I-T) = \mathcal{X}$, then by (iii) we get that $\ker(I-T^*) = \{0\}$. So, as $T^*$ is compact, $1 \notin \sigma(T^*) = \sigma(T)$ and $\ker(I-T) = \{0\}$.

(v) Since $\operatorname{ran}(I-T)$ is closed, we get from Lemma 9.7.1 that

$$\dim\ker(I-T^*) = \dim(\mathcal{X}/\operatorname{ran}(I-T))^*.$$

As $\dim\ker(I-T)^* < \infty$, we get that $(\mathcal{X}/\operatorname{ran}(I-T))^*$ is finite-dimensional, and this implies $\dim\mathcal{X}/\operatorname{ran}(I-T) < \infty$ by Lemma 5.7.8. As the dual of a finite-dimensional space has the same dimension as the space (Proposition 1.7.11), we have shown that

$$\dim\ker(I-T^*) = \dim\mathcal{X}/\operatorname{ran}(I-T).$$

(vi) Let $k = \dim\ker(I-T)$, $\ell = \dim\ker(I-T^*) = \dim\mathcal{X}/\operatorname{ran}(I-T)$. Suppose that $k < \ell$. By Proposition 5.7.27 we know that $\ker(I-T)$ and $\operatorname{ran}(I-T)$ are complemented, the latter with a complement of dimension $\ell$. Say $P$ is a bounded projection onto $\ker(I-T)$, and $M$ is a complement for $\operatorname{ran}(I-T)$, with $\dim M = \ell$. Since $k < \ell$, there exists a linear injection $\Phi : \ker(I-T) \to M$, that is not surjective. Let $S = T + \Phi P$. Since $P$ is finite-rank and $T$ is compact, $S$ is compact. If $(I-S)x = 0$, then

$$0 = x - Tx - \Phi Px = (x - Tx) - \Phi Px.$$

As $x - Tx \in \operatorname{ran}(I-T)$ and $\Phi Px \in M$, we get that $x - Tx = \Phi Px = 0$. So $x = Tx$, and $Px = 0$ since $\Phi$ is injective. Then $x \in \ker(I-T)$, implying that $x = Px = 0$. So $\ker(I-S) = \{0\}$; by (iv), $\operatorname{ran}(I-S) = \mathcal{X}$. Let $y \in M \setminus \operatorname{ran}\Phi$. Since $\operatorname{ran}(I-S) = \mathcal{X}$, there exists $x \in \mathcal{X}$ with $(I-S)x = y$. That is,

$$x - Tx - \Phi Px = y.$$

As $y \in M$ and $x - Tx \in \operatorname{ran}(I-T)$, since $M \cap \operatorname{ran}(I-T) = \{0\}$ the equality above implies that $y = -\Phi Px$, a contradiction. It follows that $\ell \leq k$, that is $\dim\ker(I-T^*) \leq \dim\ker(I-T)$.

If we now apply the above to $T^*$, we get

$$\dim\ker(I-T^{**}) \leq \dim\ker(I-T^*) \leq \dim\ker(I-T).$$

Now every $x \in \ker(I-T)$ gives us $\hat{x} \in \ker(I-T^{**})$. Indeed, for any $\varphi \in \mathcal{X}^*$,

$$((I-T^{**})\hat{x})\varphi = \hat{x}((I-T^*)\varphi) = ((I-T^*)\varphi)x = \varphi((I-T)x) = \varphi(0) = 0.$$

Thus $\dim\ker(I-T) \leq \dim\ker(I-T^{**}) \leq \dim\ker(I-T^*) \leq \dim\ker(I-T)$, and so

$$\dim\ker(I-T) = \dim\ker(I-T^*). \qquad \square.$$

**9.7.3. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. An operator $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ is* ***Fredholm*** *if $\dim\ker T < \infty$ and $\dim \mathcal{Y}/\operatorname{ran} T < \infty$. We denote the set of Fredholm operators in $\mathcal{B}(\mathcal{X},\mathcal{Y})$ by $\mathfrak{F}(\mathcal{X},\mathcal{Y})$ The* ***index*** *of $T$ is*

$$\operatorname{Ind} T = \dim\ker T - \dim\operatorname{coker} T,$$

*where the space $\operatorname{coker} T = \mathcal{Y}/\operatorname{ran} T$ is called the* ***cokernel*** *of $T$.*

For the rest of this section, $\mathcal{X}$ and $\mathcal{Y}$ will always be Banach spaces. An interesting first consequence of the definition is that the range of a Fredholm operator is closed (Proposition 9.7.4). This guarantees that the cokernel $\mathcal{Y}/\operatorname{ran} T$ is a Banach space (see page 354).

***9.7.4. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ be a Fredholm operator. Then $\operatorname{ran} T$ is closed and complemented, with finite-dimensional complement.*

*Proof.* As $\dim\ker T < \infty$, it follows that $\ker T$ is closed and complemented (Proposition 5.7.27). So there exists a closed subspace $\mathcal{X}' \subset \mathcal{X}$ such that $\mathcal{X} = \mathcal{X}' + \ker T$ as a topological direct sum. If $x \in \mathcal{X}'$ and $Tx = 0$, then $x \in \mathcal{X}' \cap \ker T = \{0\}$, so $T|_{\mathcal{X}'}$ is injective.

Since $\dim \mathcal{Y}/\operatorname{ran} T < \infty$, there is a basis $y_1 + \operatorname{ran} T, \ldots, y_m + \operatorname{ran} T$. Let $W = \operatorname{span}\{y_1, \ldots, y_m\}$. It follows that $\mathcal{Y} = W + \operatorname{ran} T$ as direct sum (a priori not topological). Now consider the Banach space $W \oplus \mathcal{X}'$ with the norm $\|(w,x)\| = \|w\| + \|x\|$, and define $\tilde{T} : W \oplus \mathcal{X}' \to \mathcal{Y}$ by $\tilde{T}(w,x) = w + Tx$. We have, with $c = \max\{\|T\|, 1\}$,

$$\|\tilde{T}(w,x)\| = \|w + Tx\| \leq \|w\| + \|T\|\,\|x\| \leq c\,(\|w\| + \|x\|) = c\,\|(w,x)\|.$$

So $\tilde{T}$ is bounded. It is also surjective (since $\mathcal{Y} = W + \operatorname{ran} T$) and injective (if $w+Tx = 0$ then $w = 0$ and $Tx = 0$; and $T$ is injective on $\mathcal{X}'$ so $x = 0$) between Banach spaces. By the Open Mapping Theorem (6.3.5), the operator $\tilde{T}$ is open. Being bijective, it also maps closed sets to closed sets (as a bijection maps complements to complements). So

$$\operatorname{ran} T = \tilde{T}(\{0\} \oplus \mathcal{X}')$$

is closed. Knowing $\operatorname{ran} T$ is closed with finite-dimensional cokernel, by Proposition 5.7.27 it is complemented with finite-dimensional complement. □

*9.7.5. Example.* Any $T \in \mathcal{B}(\mathcal{X})$, invertible, is Fredholm with $\operatorname{Ind} T = 0$. Indeed, if $T$ is invertible then $\ker T = \{0\}$ and $\operatorname{ran} T = \mathcal{X}$, so $\dim\ker T = \dim \mathcal{X}/\operatorname{ran} T = 0$. The converse is not true, though. Let $\mathcal{X} = \ell^2(\mathbb{N})$, with the canonical basis $\{e_n\}$. Fix $k \in \mathbb{N}$. Let $T \in \mathcal{B}(\mathcal{X})$ be the orthogonal projection onto $\overline{\operatorname{span}}^{\|\cdot\|}\{e_{k+1}, e_{k+2}, \ldots\}$. Then $\ker T = \operatorname{span}\{e_1, \ldots, e_k\}$, $\operatorname{ran} T = \overline{\operatorname{span}}\{e_n : n > k\}$. So $\operatorname{ran} T$ is closed, $\dim\ker T = k = \dim \mathcal{X}/\operatorname{ran} T$. Thus $T$ is Fredholm, with $\operatorname{Ind} T = 0$, and not invertible. It is invertible on "most" of $\mathcal{X}$, though; and that's precisely what characterizes Fredholm operators, as we will see in Proposition 9.7.7. Here we have $T = I - P_k$, a compact perturbation of a Fredholm operator, and that is why it is Fredholm.

*9.7.6. Example.* Let $S \in \mathcal{B}(\ell^2(\mathbb{N}))$ be the unilateral shift (cfr. Examples 9.5.10 and 9.5.11). Since $S$ is an isometry, it is injective so $\dim\ker S = 0$. The range of $S$ is $(\mathbb{C}\, e_1)^\perp$, and hence $\dim\operatorname{coker} S = 1$. So $\operatorname{Ind} S = -1$. One can similarly check that $\operatorname{Ind} S^n = -n$, and that for the left shift $T$, $\operatorname{Ind} T^n = n$.

**9.7.7. Proposition.**

*An operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is Fredholm if and only if there exists $S \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ such that $I_{\mathcal{X}} - ST \in \mathcal{K}(\mathcal{X})$, $I_{\mathcal{Y}} - TS \in \mathcal{K}(\mathcal{Y})$. In other words, $T$ is Fredholm if and only if it is invertible modulo the compact operators. In particular, $T$ is Fredholm if and only if $T + K$ is Fredholm for all $K$ compact.*

*Proof.* The finite-dimensional case is trivial, as we can take $S = 0$. So assume that $\dim \mathcal{X} = \infty$. Suppose first that $T$ is Fredholm. By Proposition 5.7.27 the finite-dimensional subspace $\ker T$ is complemented; that is, there exists a projection $P \in \mathcal{B}(\mathcal{X})$ with $P\mathcal{X} = \ker T$. And by Proposition 9.7.4, $\operatorname{ran} T$ is closed and complemented, with a finite-rank projection $Q$ onto this complement. If we consider the restriction operator $T : (I_{\mathcal{X}} - P)\mathcal{X} \to \operatorname{ran} T$, this operator is bounded and bijective between Banach spaces, so invertible by the Inverse Mapping Theorem (6.3.6). Define $S$ to be the inverse of this restriction, extended as zero on the complement of $\operatorname{ran} T$ (this can be done because $\operatorname{ran} T$ is complemented!). Given any $x \in \mathcal{X}$, we can write $x = v + w$, in a unique way, with $v = Px \in \ker T$ and $w = (I_{\mathcal{X}} - P)x$ in its complement $(I_{\mathcal{X}} - P)\mathcal{X}$. Then $STx = STw = w = (I_{\mathcal{X}} - P)x$. So $I_{\mathcal{X}} - ST = P \in \mathcal{K}(\mathcal{X})$. Given $y \in \mathcal{Y}$, we can write $y = w + Tx$, with $x \in (I_{\mathcal{X}} - P)\mathcal{X}$ and $w$ in the complement of $\operatorname{ran} T$. Then $TSy = TS(w + Tx) = TSTx = Tx = (I_{\mathcal{Y}} - Q)y$. Thus $I_{\mathcal{Y}} - TS = Q \in \mathcal{K}(\mathcal{X})$.

Conversely, suppose that $I_{\mathcal{X}} - ST$ and $I_{\mathcal{Y}} - TS$ are compact. By (i) in Proposition 9.7.2 we get $\dim \ker ST < \infty$, so

$$\dim \ker T \leq \dim \ker ST < \infty.$$

We also get from Proposition 9.7.2 that $\dim \mathcal{X}/\operatorname{ran} TS < \infty$. As $\operatorname{ran} TS \subset \operatorname{ran} T$,

$$\dim \mathcal{X}/\operatorname{ran} T \leq \dim \mathcal{X}/\operatorname{ran} TS < \infty,$$

since there is a canonical linear surjection $\mathcal{X}/\operatorname{ran} TS \to \mathcal{X}/\operatorname{ran} T$. So $T$ is Fredholm. □

An unexpected consequence of the proof of Proposition 9.7.7 is that an operator is Fredholm if and only if it is invertible modulo the finite-rank operators. That is, if there exists $S$ such that $I - ST$ and $I - TS$ are compact, then there exists $S'$ such that $I - S'T$ and $I - TS'$ are finite-rank.

It also follows from Proposition 9.7.7 that when a composition of Fredholm operators makes sense, the composition is Fredholm (Exercise 9.7.4).

Here is a consequence that may help us see what the index is measuring.

***9.7.8. Corollary.***

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. The following statements are equivalent:*

*(i)* *$T$ is Fredholm;*

*(ii)* *there exist decompositions $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$, $\mathcal{Y} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$, where*
- *(a)* *$\mathcal{X}_1, \mathcal{Y}_1$ are closed;*
- *(b)* *$\dim \mathcal{X}_2 < \infty$, $\dim \mathcal{Y}_2 < \infty$;*

*(c) with respect to this decomposition,*

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix},$$

*where* $T_{11} \in \mathcal{B}(\mathcal{X}_1, \mathcal{Y}_1)$ *is invertible.*

*In this situation, when* $T$ *is Fredholm,* $\operatorname{Ind} T = \dim \mathcal{X}_2 - \dim \mathcal{Y}_2$.

*Proof.* (i) $\Longrightarrow$ (ii) we have that $T$ is Fredholm. Define $\mathcal{X}_2 = \ker T$. Since $\dim \mathcal{X}_2 < \infty$, it is complemented by Proposition 5.7.27. Let $\mathcal{X}_1$ be such a complement. Let $\mathcal{Y}_1 = \operatorname{ran} T$; it is closed by Proposition 9.7.4. Again by Proposition 5.7.27 we have that $\mathcal{Y}_1$ is complemented with a complement $\mathcal{Y}_2$ with $\dim \mathcal{Y}_2 = \dim \mathcal{Y}/\mathcal{Y}_1$. This satisfies the first two conditions automatically. Because $\mathcal{X}_2 = \ker T$, $T\mathcal{X}_1 = \operatorname{ran} T = \mathcal{Y}_1$. Then $T_{11} = T|_{\mathcal{X}_1} : \mathcal{X}_1 \to \mathcal{Y}_1$ is a bijective bounded operator (Exercise 9.7.1), hence invertible by the Inverse Mapping Theorem (6.3.6). Then

$$T = \begin{bmatrix} T_{11} & 0 \\ 0 & 0 \end{bmatrix}.$$

And, by construction, $\operatorname{Ind} T = \dim \ker T - \dim \operatorname{coker} T = \dim \mathcal{X}_2 - \dim \mathcal{Y}_2$.

(ii) $\Longrightarrow$ (i) let $S : \mathcal{Y} \to \mathcal{X}$ be

$$S = \begin{bmatrix} T_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

We have

$$ST = \begin{bmatrix} T_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} I_{\mathcal{X}_1} & T_{11}^{-1}T_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_{\mathcal{X}_1} & 0 \\ 0 & I_{\mathcal{X}_2} \end{bmatrix} + \begin{bmatrix} 0 & T_{11}^{-1}T_{12} \\ 0 & -I_{\mathcal{X}_2} \end{bmatrix}.$$

Since $T_{12} : \mathcal{X}_2 \to \mathcal{Y}_1$ has finite-dimensional domain, it is finite-rank, hence compact. And $-I_{\mathcal{X}_2} : \mathcal{X}_2 \to \mathcal{Y}_2$ is also finite-rank. So the last $2 \times 2$ matrix above is finite-rank, and this shows that $ST - I_{\mathcal{X}} \in \mathcal{K}(\mathcal{X})$. With the exact same idea one shows that $TS - I_{\mathcal{Y}} \in \mathcal{K}(\mathcal{Y})$. By Proposition 9.7.7, $T$ is Fredholm.

Now we need to calculate the index. In this block-matrix presentation, $\ker T$ consists of those $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $T_{11}x_1 + T_{12}x_2 = 0$ and $T_{21}x_1 + T_{22}x_2 = 0$. Using that $T_{11}$ is invertible, the first equality is equivalent to

$$x_1 = -T_{11}^{-1}T_{12}x_2 \tag{9.35}$$

and, substituting this into the second equation,

$$(T_{22} - T_{21}T_{11}^{-1}T_{12})x_2 = 0. \tag{9.36}$$

Let $R : \mathcal{X}_2 \to \mathcal{Y}_2$ be given by $R = T_{22} - T_{21}T_{11}^{-1}T_{12}$. We will show that $\ker R \simeq \ker T$. Namely, consider $Q : \ker R \to \ker T$ given by

$$Qx = \begin{bmatrix} -T_{11}^{-1}T_{12}x \\ x \end{bmatrix}.$$

The conditions (9.35) and (9.36) imply $\operatorname{ran} Q = \ker T$ (Exercise 9.7.2). And $Q$ is trivially injective, so it is an isomorphism that gives $\dim\ker T = \dim\ker R$. Now define $P : \mathcal{Y}/\operatorname{ran} T \to \mathcal{Y}_2/\operatorname{ran} R$ by

$$P : \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \operatorname{ran} T \longmapsto y_2 - T_{21}T_{11}^{-1}y_1 + \operatorname{ran} R.$$

Then $P$ is linear, and we need to check that $P$ is well-defined, injective, and surjective. If $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \operatorname{ran} T$, then

$$y_1 = T_{11}x_1 + T_{12}x_2, \qquad\qquad y_2 = T_{21}x_1 + T_{22}x_2,$$

for some $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. Then

$$\begin{aligned} y_2 - T_{21}T_{11}^{-1}y_1 &= T_{21}x_1 + T_{22}x_2 - T_{21}T_{11}^{-1}(T_{11}x_1 + T_{12}x_2) \\ &= T_{22}x_2 - T_{21}T_{11}^{-1}T_{12}x_2 = Rx_2 \in \operatorname{ran} R \end{aligned}$$

and $P$ is well-defined. To see that $P$ is injective, if $y_2 - T_{21}T_{11}^{-1}y_1 \in \operatorname{ran} R$ then there exists $x_2 \in \mathcal{X}_2$ with

$$y_2 - T_{21}T_{11}^{-1}y_1 = T_{22}x_2 - T_{21}T_{11}^{-1}T_{12}x_2.$$

We can write this as

$$y_2 = T_{21}[T_{11}^{-1}(y_1 - T_{12}x_2)] + T_{22}x_2.$$

If we put $x_1 = T_{11}^{-1}(y_1 - T_{12}x_2)$, then the above equality is $y_2 = T_{21}x_1 + T_{22}x_2$; and we have

$$T_{11}x_1 + T_{12}x_2 = y_1 - T_{12}x_2 + T_{12}x_2 = y_1.$$

Thus $y \in \operatorname{ran} T$ and $P$ is injective.

To check that $P$ is surjective, given any $y_2 \in \mathcal{Y}_2$ we have

$$y_2 + \operatorname{ran} R = P\left( \begin{bmatrix} 0 \\ y_2 \end{bmatrix} + \operatorname{ran} T \right).$$

The isomorphisms $P$ and $Q$ show that

$$\operatorname{Ind} T = \dim\ker T - \dim\operatorname{coker} T = \dim\ker R - \dim\operatorname{coker} R = \operatorname{Ind} R.$$

Now $R$ is a linear operator between finite-dimensional spaces. So we have, by the rank-nullity theorem (Exercise 5.2.6),

$$\dim \mathcal{X}_2 = \dim\ker R + \dim\operatorname{ran} R = \dim\ker R + \dim \mathcal{Y}_2 - \dim\operatorname{coker} R.$$

Hence

$$\operatorname{Ind} T = \operatorname{Ind} R = \dim \mathcal{X}_2 - \dim \mathcal{Y}_2. \qquad \square$$

**9.7.9. Corollary.**

*Let $\mathcal{X},\mathcal{Y}$ be Banach spaces. Then $\mathfrak{F}(\mathcal{X},\mathcal{Y})$ is open, and the index map* $\operatorname{Ind}: T \longmapsto \operatorname{Ind} T$ *is constant on each connected component of $\mathfrak{F}(\mathcal{X},\mathcal{Y})$.*

*Proof.* Let $T \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ be Fredholm. Let $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, T_{11}, T_{12}, T_{21}, T_{22}$ as in Corollary 9.7.8. Let $S \in \mathcal{B}(\mathcal{X},\mathcal{Y})$ such that

$$\|S - T\| \leq \frac{1}{1 + \|P_1\|\,\|T_{11}^{-1} \oplus I_{\mathcal{X}_2}\|},$$

where $P_1 = I - P_2$ is the projection onto $\mathcal{X}_1$. Expressing $S$ in terms of the decompositions $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ and $\mathcal{Y} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$,

$$\begin{aligned}
\|I_{\mathcal{X}_1} - T_{11}^{-1}S_{11}\| &= \left\| P_1 \left( \begin{bmatrix} T_{11}^{-1}S_{11} & T_{11}^{-1}S_{12} \\ S_{21} & S_{22} \end{bmatrix} - \begin{bmatrix} I_{\mathcal{X}_1} & T_{11}^{-1}T_{12} \\ T_{21} & T_{22} \end{bmatrix} \right) P_1 \right\| \\
&\leq \|P_1\| \left\| \begin{bmatrix} T_{11}^{-1}S_{11} & T_{11}^{-1}S_{12} \\ S_{21} & S_{22} \end{bmatrix} - \begin{bmatrix} I_{\mathcal{X}_1} & T_{11}^{-1}T_{12} \\ T_{21} & T_{22} \end{bmatrix} \right\| \\
&= \|P_1\| \left\| \begin{bmatrix} T_{11}^{-1} & 0 \\ 0 & I_{\mathcal{X}_2} \end{bmatrix} \left( \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} - \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \right) \right\| \\
&\leq \|P_1\|\,\|T_{11}^{-1} \oplus I_{\mathcal{X}_2}\|\,\|S - T\| \\
&< 1.
\end{aligned}$$

Then $T_{11}^{-1}S_{11}$ is invertible by Proposition 6.2.3, so $S_{11}$ is invertible. By Corollary 9.7.8, $S$ is Fredholm and $\operatorname{Ind} S = \dim \mathcal{X}_2 - \dim \mathcal{Y}_2 = \operatorname{Ind} T$. We have shown that every operator in a small ball around $T$ is Fredholm and the index remains constant; by Proposition 1.8.15, the index is constant on each connected component. ☐

When $\mathcal{Y} = \mathcal{X}$, $\mathcal{B}(\mathcal{X})$ is a Banach algebra and the fact that $\mathfrak{F}(\mathcal{X})$ is open can be obtained rather directly by noticing that $\mathfrak{F}(\mathcal{X}) = \pi^{-1}(GL(\mathcal{B}(\mathcal{X})/\mathcal{K}(\mathcal{X})))$, and this latter set is open as it is a continuous preimage of an open set (the relevant results being Propositions 9.2.2 and 5.3.13). The argument for the continuity of the index still requires some more work, which makes the proof of Corollary 9.7.9 look fairly slick.

Another consequence of Proposition 9.7.7 is that it allows us to calculate the index of the adjoint.

**9.7.10. Corollary.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Then $T$ is Fredholm if and only if $T^*$ is Fredholm, and in that case*

$$\operatorname{Ind} T^* = -\operatorname{Ind} T.$$

*Proof.* Suppose that $T$ is Fredholm. By Proposition 9.7.7 there exists $S \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ with $ST - I_{\mathcal{X}}$ and $TS - I_{\mathcal{Y}}$ compact. By Theorem 9.6.12, $T^*S^* - I_{\mathcal{X}^*}$ and $S^*T^* - I_{\mathcal{Y}^*}$ are compact, and so $T^*$ is Fredholm.

Conversely, suppose that $T^*$ is Fredholm. Let $\Gamma : \ker T \to (\mathcal{X}^*/\operatorname{ran} T^*)^*$ be

$$(\Gamma x)(\varphi + \operatorname{ran} T^*) = \varphi(x).$$

This is well-defined: if $\varphi - \varphi' \in \operatorname{ran} T^*$, then $\varphi - \varphi' = T^*\psi$ for some $\psi \in \mathcal{Y}^*$. Then, for any $x \in \ker T$,

$$(\varphi - \varphi')(x) = (T^*\psi)x = \psi(Tx) = 0,$$

so $\Gamma$ is well-defined. It is obviously linear. If $\Gamma(x) = 0$, then $\varphi(x) = 0$ for all $\varphi \in \mathcal{X}^*$, so $x = 0$ by Hahn–Banach (Corollary 5.7.7) and $\Gamma$ is injective. Now let $\psi \in (\mathcal{X}^*/\operatorname{ran} T^*)^*$. Define $\tilde{\psi} : \mathcal{X}^* \to \mathbb{C}$ by $\tilde{\psi}(\varphi) = \psi(\varphi + \operatorname{ran} T^*)$. As $\mathcal{X}^*/\operatorname{ran} T^*$ is finite-dimensional, $\psi$ is bounded. Then

$$|\tilde{\psi}(\varphi)| = |\psi(\varphi + \operatorname{ran} T^*)| \leq \|\psi\| \, \|\varphi + \operatorname{ran} T^*\| \leq \|\psi\| \, \|\varphi\|.$$

Thus $\tilde{\psi} \in \mathcal{X}^{**}$. If $\varphi_j \xrightarrow{\text{weak}^*} 0$, this means that $\varphi_j(x) \to 0$ for all $x \in \mathcal{X}$. Again because $\mathcal{X}^*/\operatorname{ran} T^*$ is finite-dimensional, all vector space topologies agree (Theorem 5.4.16). In particular, the weak*-quotient topology (see Lemma 5.4.20) agrees with the norm topology. Thus $\varphi_j + \operatorname{ran} T^* \to 0$, and then

$$\tilde{\psi}(\varphi_j) = \psi(\varphi_j + \operatorname{ran} T^*) \to 0,$$

which means that $\tilde{\psi}$ is weak*-continuous. By Proposition 7.2.10, there exists $x_0 \in \mathcal{X}$ with $\tilde{\psi}(\varphi) = \varphi(x_0)$. In particular, $\Gamma x_0 = \psi$. For any $\rho \in \mathcal{Y}^*$,

$$\rho(Tx_0) = (T^*\rho)x_0 = \tilde{\psi}(T^*\rho) = \psi(T^*\rho + \operatorname{ran} T^*) = \psi(\operatorname{ran} T^*) = 0.$$

As this can be done for any $\rho$, we obtain (via Corollary 5.7.7) that $Tx_0 = 0$, i.e. $x_0 \in \ker T$. So $\Gamma$ is surjective, and then

$$\dim \ker T = \dim(\mathcal{X}^*/\operatorname{ran} T^*)^* = \dim \mathcal{X}^*/\operatorname{ran} T^* \tag{9.37}$$

Note that $\operatorname{ran} T$ is closed by (vii) in Proposition 9.4.6, so by Lemma 9.7.1

$$\dim \ker T^* = \dim(\mathcal{Y}/\operatorname{ran} T)^* = \dim \mathcal{Y}/\operatorname{ran} T, \tag{9.38}$$

where the second equality comes from the fact that $\dim \ker T^* < \infty$ and by the first equality the dual $(\mathcal{Y}/\operatorname{ran} T)^*$ is finite-dimensional, which forces $\mathcal{Y}/\operatorname{ran} T$ to be finite-dimensional by Lemma 5.7.8, and then Exercise 5.5.3 justifies the equality.

Combining (9.37) with (9.38) we get that $T$ is Fredholm and that
$$\operatorname{Ind} T = -\operatorname{Ind} T^*.$$ □

It is tempting to go for a short argument in the second part of the proof of Corollary 9.7.10, similar to the first part. One can obtain, from Proposition 9.7.7, an operator $S' \in \mathcal{B}(\mathcal{X}^*, \mathcal{Y}^*)$ with $I_{\mathcal{Y}^*} - S'T^*$ and $I_{\mathcal{X}^*} - T^*S'$ compact. The problem is that there is no obvious way to say that $S'$ is an adjoint, as adjoints are precisely the weak*-continuous operators (see Proposition 9.4.2 and Exercise 9.4.10).

**9.7.11. Definition.**

*Let $A_1, \ldots, A_n$ be vector spaces with morphisms $\gamma_k : A_k \to A_{k+1}$, $k = 1, \ldots, n-1$. The sequence*
$$A_1 \xrightarrow{\gamma_1} A_2 \xrightarrow{\gamma_2} A_3 \xrightarrow{\gamma_3} \cdots \xrightarrow{\gamma_{n-1}} A_n$$
*is an* ***exact sequence*** *if $\ker \gamma_{k+1} = \operatorname{ran} \gamma_k$ for $k = 1, \ldots, n-1$.*

Exact sequences play a big role in large parts of algebra. The most common case is that of a **short exact sequence**
$$0 \to \mathcal{X} \to \mathcal{Y} \to \mathcal{Z} \to 0.$$
Note that there is no ambiguity in what $\gamma_1$ and $\gamma_4$ should be. In any short sequence we have that $\mathcal{Z} \simeq \mathcal{Y}/\mathcal{X}$.

We will need the following result:

***9.7.12. Lemma.***

*Let*
$$0 \xrightarrow{\gamma_0} \mathcal{X}_1 \xrightarrow{\gamma_1} \mathcal{X}_2 \xrightarrow{\gamma_2} \mathcal{X}_3 \xrightarrow{\gamma_3} \cdots \xrightarrow{\gamma_{n-1}} \mathcal{X}_n \xrightarrow{\gamma_n} 0$$
*be an exact sequence of finite-dimensional vector spaces. Then*
$$\sum_{k=1}^{n} (-1)^k \dim \mathcal{X}_k = 0.$$

*Proof.* We proceed by induction. Consider first the sequence
$$0 \to \mathcal{X}_1 \xrightarrow{\gamma_1} \operatorname{ran} \gamma_1 \longrightarrow 0.$$
We have that $\gamma_1$ is injective because of the leftmost arrow, so
$$\dim \mathcal{X}_1 - \dim \operatorname{ran} \gamma_1 = 0.$$
Now fix $k < n$ and assume that the exact sequence
$$0 \xrightarrow{\gamma_0} \mathcal{X}_1 \xrightarrow{\gamma_1} \mathcal{X}_2 \xrightarrow{\gamma_2} \mathcal{X}_3 \xrightarrow{\gamma_3} \cdots \xrightarrow{\gamma_{k-1}} \mathcal{X}_k \xrightarrow{\gamma_k} \operatorname{ran} \gamma_k \longrightarrow 0.$$

satisfies $\sum_{j=1}^{k}(-1)^j \dim \mathcal{X}_j = (-1)^k \dim \operatorname{ran} \gamma_k$. For

$$0 \xrightarrow{\gamma_0} \mathcal{X}_1 \xrightarrow{\gamma_1} \mathcal{X}_2 \xrightarrow{\gamma_2} \mathcal{X}_3 \xrightarrow{\gamma_3} \cdots \xrightarrow{\gamma_k} \mathcal{X}_{k+1} \xrightarrow{\gamma_{k+1}} \operatorname{ran} \gamma_{k+1} \longrightarrow 0$$

we have, since $\gamma_{k+1}$ is surjective and $\dim \ker \gamma_{k+1} = \dim \operatorname{ran} \gamma_k$,

$$\dim \operatorname{ran} \gamma_{k+1} = \dim \mathcal{X}_{k+1} - \dim \ker \gamma_{k+1} = \dim \mathcal{X}_{k+1} - \dim \operatorname{ran} \gamma_k.$$

Then

$$\begin{aligned}\sum_{j=1}^{k+1}(-1)^j \dim \mathcal{X}_j &= (-1)^k \dim \operatorname{ran} \gamma_k + (-1)^{k+1} \dim \mathcal{X}_{k+1} \\ &= (-1)^k \dim \operatorname{ran} \gamma_k + (-1)^{k+1}(\dim \operatorname{ran} \gamma_k + \dim \operatorname{ran} \gamma_{k+1}) \\ &= (-1)^{k+1} \dim \operatorname{ran} \gamma_{k+1}.\end{aligned}$$

This completes the induction and so

$$\sum_{k=1}^{n}(-1)^k \dim \mathcal{X}_k = (-1)^n \dim \operatorname{ran} \gamma_n = 0. \qquad \square$$

A trivial application of the above result is the following. Let $\mathcal{X}$ be a finite-dimensional vector space and $\mathcal{Y} \subset \mathcal{X}$. Then we have an exact sequence

$$0 \to \mathcal{Y} \to \mathcal{X} \to \mathcal{X}/\mathcal{Y} \to 0.$$

By Lemma 9.7.12,

$$\dim \mathcal{X}/\mathcal{Y} = \dim \mathcal{X} - \dim \mathcal{Y}.$$

Of course, one usually proves this relation directly.

The application of Lemma 9.7.12 in the context of Fredholm operators is the following.

**9.7.13. *Proposition.***

*Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be Fredholm operators. Then*

$$\operatorname{Ind} ST = \operatorname{Ind} S + \operatorname{Ind} T.$$

*Proof.* We have an exact sequence

$$0 \longrightarrow \ker T \xrightarrow{i} \ker ST \xrightarrow{T} \ker S \xrightarrow{q} \operatorname{coker} T \xrightarrow{\tilde{S}} \operatorname{coker} ST \xrightarrow{j} \operatorname{coker} S \to 0,$$

where $i$ is the identity, $q$ is the quotient map, $\tilde{S}$ is the map $y + \operatorname{ran} T \longmapsto Sy + \operatorname{ran} ST$, and $j$ is the map $z + \operatorname{ran} ST \longmapsto z + \operatorname{ran} S$. The sequence is exact because

$$\operatorname{ran} i = \ker T, \qquad \ker q = \ker S \cap \operatorname{ran} T = \operatorname{ran} T|_{\ker ST},$$

$$\ker j = \{z + \operatorname{ran} ST : \ z \in \operatorname{ran} S\} = \operatorname{ran} \tilde{S},$$

and $j$ is obviously surjective. The only less obvious exactness is $\operatorname{ran} q = \ker \tilde{S}$. We have
$$\operatorname{ran} q = \{y + \operatorname{ran} T : \ y \in \ker S\} \subset \ker \tilde{S}.$$
Conversely, if $y + \operatorname{ran} T \in \ker \tilde{S}$, then there exists $x$ such that $Sy = STx$. Then $y - Tx \in \ker S$, so $y \in \ker S + \operatorname{ran} T = \ker q$, which shows that $\ker \tilde{S} \subset \operatorname{ran} q$.

Having established the exactness in the sequence, Lemma 9.7.12 gives us that
$$\begin{aligned}\dim\ker T - \dim\ker ST + \dim\ker S - \dim\operatorname{coker} T \\ + \dim\operatorname{coker} ST - \dim\operatorname{coker} S = 0.\end{aligned}$$
Rearranged, this is
$$\begin{aligned}\dim\ker ST - \dim\operatorname{coker} ST = \dim\ker T - \dim\operatorname{coker} T \\ + \dim\ker S - \dim\operatorname{coker} S.\end{aligned}$$
This is precisely
$$\operatorname{Ind} ST = \operatorname{Ind} S + \operatorname{Ind} T. \qquad \square$$

**9.7.14. Corollary.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$, $K \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$. If $T$ is Fredholm, then*
$$\operatorname{Ind}(T + K) = \operatorname{Ind} T.$$

*Proof.* Since $T$ is Fredholm, by Proposition 9.7.7 there exist $S \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ and $K' \in \mathcal{K}(\mathcal{X})$ with $ST = I_{\mathcal{X}} + K'$. We have, by (v) and (vi) in Proposition 9.7.2, that $\operatorname{Ind}(ST) = \operatorname{Ind}(I_{\mathcal{X}} + K') = 0$. By Proposition 9.7.13
$$\operatorname{Ind} S + \operatorname{Ind} T = \operatorname{Ind}(ST) = \operatorname{Ind}(I + K) = 0.$$
As $S(T + K) = I_{\mathcal{X}} + (K' + SK)$,
$$\operatorname{Ind} S + \operatorname{Ind}(T + K) = \operatorname{Ind}(I_{\mathcal{X}} + (K' + SK)) = 0.$$
Then $\operatorname{Ind}(T + K) = -\operatorname{Ind}(S) = \operatorname{Ind} T$. $\square$

A particular consequence of Corollary 9.7.14 is that
$$\operatorname{Ind}(\lambda I + K) = 0, \qquad K \in \mathcal{K}(\mathcal{X}),\ \lambda \in \mathbb{C} \setminus \{0\}.$$

**9.7.15. Definition.**

*Let $T \in \mathcal{B}(\mathcal{X})$. The* ***essential spectrum of*** $T$ *is the set*
$$\sigma_{\text{ess}}(T) = \{\lambda \in \mathbb{C} : \ T - \lambda I \ \text{ is not Fredholm}\}.$$

***9.7.16. Corollary.***

*Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$. Then $\sigma_{\text{ess}}(T) = \sigma_{\text{ess}}(T^*)$, and $\sigma_{\text{ess}}(T+K) = \sigma_{\text{ess}}(T)$ for all compact $K$.*

*Proof.* By Corollary 9.7.10, $T - \lambda I_{\mathcal{X}}$ is Fredholm if and only if $T^* - \lambda I_{\mathcal{X}^*}$ is Fredholm; so $\sigma_{\text{ess}}(T) = \sigma_{\text{ess}}(T^*)$. And by Proposition 9.7.7 $T + K - \lambda I_{\mathcal{X}}$ is Fredholm if and only if $T - \lambda I_{\mathcal{X}}$ is Fredholm, showing that $\sigma_{\text{ess}}(T+K) = \sigma_{\text{ess}}(T)$. □

*9.7.17. Example.* For the unilateral shift $S \in \mathcal{B}(\ell^2(\mathbb{N}))$ we studied its spectrum in Example 9.5.10. Since $S^*S = I$ and $I - SS^*$ is compact, on the Calkin algebra $S$ becomes a unitary. Thus $\sigma_{\text{ess}}(S) \subset \mathbb{T}$. We know that $\sigma(S) = \overline{\mathbb{D}}$, $\sigma_p(S) = \varnothing$, and $\sigma_{ap}(S) = \mathbb{T}$. Thus for $|\lambda| = 1$ we have that $S - \lambda I$ is not bounded below and $\ker(S - \lambda I) = \{0\}$, forcing $\operatorname{ran}(S - \lambda I)$ to be not closed by Proposition 9.5.4. Therefore $S - \lambda I$ is not Fredholm, and $\lambda \in \sigma_{\text{ess}}(S)$. We did this for all $\lambda \in \mathbb{T}$, so $\sigma_{\text{ess}}(S) = \mathbb{T}$. By Corollary 9.7.10 we have that $\sigma_{\text{ess}}(S^*) = \mathbb{T}$.

*9.7.18. Example.* Consider the bilateral shift $V \in \mathcal{B}(\ell^p(\mathbb{Z}))$. We know from Example 9.5.16 that $\sigma(V) = \mathbb{T}$. Let $K = -E_{1,0}$. That is, $Ke_0 = -e_1$, $Ke_n = 0$ for all $n \neq 0$. Then

$$(V+K)e_n = \begin{cases} e_{n+1}, & n \neq 0 \\ 0, & n = 0 \end{cases}$$

Visually, what we are doing is thinking of $V$ as an infinite matrix with the lower subdiagonal with all entries equal to 1, and replacing one of these 1 with a 0. Consider the decomposition $\ell^p(\mathbb{Z}) = H_0 + H_1$, where $H_0, H_1 \subset \ell^p(\mathbb{Z})$ are the subspaces

$$H_0 = \overline{\operatorname{span}}\{e_n : \ n \leq 0\}, \qquad H_1 = \overline{\operatorname{span}}\{e_n : \ n \geq 1\}.$$

Both are invariant for $V + K$, and $H_0 \cap H_1 = \{0\}$. So $V + K$ can be considered as the direct sum of the two operators $S_0 = (T+K)|_{H_0}$ and $S_1 = (T+K)|_{H_1}$. The spectrum of a direct sum is the union of the spectra (Exercise 9.5.7). And $S_1$ is a unilateral shift on $H_1$, while $S_0$ is the adjoint of a unilateral shift on $H_0$. By Example 9.5.10 we know that $\sigma(S_0) = \sigma(S_1) = \overline{\mathbb{D}}$. So

$$\sigma(V+K) = \overline{\mathbb{D}},$$

showing that this rank-one perturbation of the bilateral shift changes the spectrum from $\mathbb{T}$ to $\overline{\mathbb{D}}$. By Corollary 9.7.16

$$\sigma_{\text{ess}}(V) = \sigma_{\text{ess}}(V + K).$$

By Exercise 9.7.5, $V + K - \lambda I$ is Fredholm if and only if both $S_0 - \lambda I$ and $S_1 - \lambda I$ are Fredholm. It follows that

$$\sigma_{\text{ess}}(V) = \sigma_{\text{ess}}(V + K) = \sigma_{\text{ess}}(S_0) \cup \sigma_{\text{ess}}(S_1) = \mathbb{T}.$$

*9.7.19. Example.* Given any $C \subset \mathbb{C}$ compact, there exists $T \in \mathcal{B}(\ell^2(\mathbb{N}))$ with $\sigma_{\text{ess}}(T) = C$. Indeed, if we choose a dense subset of $C$, enumerate it as $\{q_n\}$, and define $Te_n = q_n e_n$, extended by linearity and continuity, this defines a bounded $T$ with $\sigma(T) = C$ (Example 9.5.6). But it this is not enough for $\sigma_{\text{ess}}(T) = C$. As defined, $\sigma_{\text{ess}}(T)$ consists of the accumulation points in $C$. We can solve this problem in the following way. Let $\{N_k\}$ be a countable family of countable infinite subsets of $\mathbb{N}$ with $\mathbb{N} = \bigcup_k N_k$. Define a new $T$ such that for $e_n \in N_k$, $Te_n = q_k e_n$. Then $\dim\ker(T - q_n I) = \infty$ for all $n$, and so $\sigma_{\text{ess}}(T) = C$.

### 9.7.1. Exercises.

(9.7.1) In the proof of Corollary 9.7.8, show that $T_{11} : \mathcal{X}_1 \to \mathcal{Y}_1$ is a bounded linear bijection.

(9.7.2) In the proof of Corollary 9.7.8, show that $\operatorname{ran} Q = \ker T$.

(9.7.3) Let $\mathcal{X}$ be a Banach space, $T \in \mathcal{B}(\mathcal{X})$ Fredholm, $R \in \mathcal{B}(\mathcal{X})$ invertible. Show that $RT$ and $TR$ are Fredholm.

(9.7.4) Let $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ and $Z \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$ be Fredholm. Show that $ZT$ is Fredholm.

(9.7.5) Let $\mathcal{X}$ be a Banach space, and $\mathcal{X}_0, \mathcal{X}_1$ subspaces with $\mathcal{X}_0 \cap \mathcal{X}_1 = \varnothing$ and $\mathcal{X} = \mathcal{X}_0 + \mathcal{X}_1$ (that is, $\mathcal{X}$ is the direct sum of $\mathcal{X}_0$ and $\mathcal{X}_1$). Let $T \in \mathcal{B}(\mathcal{X})$ such that $T\mathcal{X}_0 \subset \mathcal{X}_0$, $T\mathcal{X}_1 \subset \mathcal{X}_1$. Let $T_0 = T|_{\mathcal{X}_0}$, $T_1 = T|_{\mathcal{X}_1}$. Show that $T$ is Fredholm if and only if $T_0$ and $T_1$ are Fredholm, and that

$$\operatorname{Ind} T = \operatorname{Ind} T_0 + \operatorname{Ind} T_1.$$

(9.7.6) Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(\mathcal{X})$ Fredholm. Show that the following statements are equivalent:

(i) there exist $S, K \in \mathcal{B}(\mathcal{X})$, with $S$ invertible and $K$ compact, such that $T = S + K$;

(ii) $\operatorname{Ind} T = 0$.

## 9.8. Schauder Bases and Basic Sequences

*This section and the next two are partly about Banach spaces, so in principle one would say those parts belong in Chapter 5. But the topics are at a different level than Chapter 5 and at some point they do use information about linear operators, and eventually obtain information about operators. Hence the location in this chapter.* Sections 9.8 and 9.9 do not play a role in the rest of the book; they should be seen as a stand–alone peek into the subtlety of Banach space techniques. Significant parts of the material in this section were directly inspired by [**AK06**, **FHH+11**]; we strongly recommend those two texts to anyone interested in a deeper understanding of Banach spaces.

One of the big selling points of Hilbert spaces is the existence of orthonormal bases; many things about Hilbert spaces are proven using them. In general, there is no analog thing in a Banach space; but at the same time the canonical basis that works in $\ell^2(\mathbb{N})$ also works as an "infinite sum basis" in $\ell^p(\mathbb{N})$ for all $1 \leq p < \infty$, and also in $c_0$. These bases are not bases in the usual algebraic sense, since a Hamel basis for a Banach space cannot be countable (Exercise 6.3.3). In finite dimension bases are just bases and always exist, so our discussion is limited to infinite-dimensional spaces.

**9.8.1. Schauder Bases.** Though we briefly discussed Schauder bases in Section 5.1, we restate the definition and we will define several related notions.

**9.8.1. Definition.**

*Let $\mathcal{X}$ be a normed space with $\dim \mathcal{X} = \infty$. A* ***Schauder basis*** *for $\mathcal{X}$ is a sequence $E = \{e_n\} \subset \mathcal{X}$ such that every $x \in \mathcal{X}$ can be written in a unique way as a series $x = \sum_n c_n e_n$ for coefficients $c_n \in \mathbb{C}$.*

*The* ***basis projections*** *of the Schauder basis $\{e_n\}$ are the maps $P_n : \mathcal{X} \to \mathcal{X}$ given by*

$$P_n : \sum_{k=1}^{\infty} c_k e_k \longmapsto \sum_{k=1}^{n} c_n e_n.$$

*The* ***basis constant*** *of $E = \{e_n\}$ is $\mathrm{b}_E = \sup_n \|P_n\|$.*

*Given a Schauder basis $\{e_n\}$ of $\mathcal{X}$ the* ***biothogonal functionals*** *$\{e_n^*\} \subset \mathcal{X}^*$ are the maps*

$$e_n^* : \sum_{k=1}^{\infty} c_n e_n \longmapsto c_n.$$

*The Schauder basis $\{e_n\}$ is* ***normalized*** *if $\|e_n\| = 1$ for all $n$, and* ***monotone*** *if $\|P_n\| = 1$ for all $n$.*

*A sequence $\{w_n\} \subset \mathcal{X}$ is a* ***basic sequence*** *if it is a Schauder basis of $\overline{\operatorname{span}}^{\|\cdot\|}\{w_n\}$. Two basic sequences $\{e_n\}$ and $\{f_n\}$ are* ***equivalent*** *if for all sequences $\{c_n\} \subset \mathbb{C}^{\mathbb{N}}$ we have that $\sum_n c_n e_n$ converges if and only $\sum_n c_n f_n$ converges.*

The uniqueness in the definition guarantees that any Schauder basis is linearly independent. And the requirement that a Schauder basis is a sequence forces $\mathcal{X}$ to be separable. There exist separable Banach spaces without Schauder basis [Enf73]; this was a nontrivial question that remained open for a long time.

Other immediate consequences of the definition are that $P_n$ is linear and bounded for all $n$, since it is finite-rank (Exercise 9.8.1). And this guarantees that $e_n^*$ is bounded, for $|e_n^*(x)| = \|e_n\|^{-1}\,\|(P_n - P_{n-1})x\|$.

A less obvious consequence is that $\{P_n\}$ is uniformly bounded, which is to say that $b_E < \infty$ for any Schauder basis (Exercise 9.8.2).

***9.8.2. Proposition.***

*Let $\mathcal{X}$ be a Banach space with Schauder basis $E = \{e_n\}$. Define*

$$\|x\|_b = \sup\{\|P_n x\| :\ n\}.$$

*Then*

*(i) $\|\cdot\|_b$ is a norm on $\mathcal{X}$, equivalent to the original norm.*

*(ii) $\|e_n^*\| \le 2\|e_n\|^{-1}\,\mathrm{b}_E$. In particular, $e_n^* \in \mathcal{X}^*$.*

*(iii) $\{e_n\}$ is a Schauder basis of $(\mathcal{X}, \|\cdot\|_b)$ with basis constant 1.*

*(iv) $1 \le \mathrm{b}_E < \infty$.*

*Proof.*

(i) A short, direct computation confirms that $\|\cdot\|_b$ is a seminorm (Exercise 9.8.3). By continuity of the norm we have

$$\|x\| = \lim_n \|P_n x\| \le \|x\|_b.$$

In particular, $\|\cdot\|_b$ is a norm. We also have $\|x\|_b \le b_E\|x\|$, so $\|\cdot\|_b$ and $\|\cdot\|$ are equivalent.

(ii) As mentioned, from

$$\|P_m x - P_{m-1}x\| = \|e_m^*(x)e_m\| \le \|e_m\|\,|e_m^*(x)|$$

we get

$$\|e_m^*\| = \sup\{|e_n^*(x)| :\ \|x\| = 1\} \le \|e_m\|^{-1}\,\|P_m - P_{m-1}\| \le 2\|e_m\|^{-1}\,\mathrm{b}_E\,.$$

(iii) With $x = \sum_n c_n e_n$,

$$\|x - P_m(x)\|_b = \sup_{n>m} \left\| \sum_{k=m+1}^{n} c_k e_k \right\| \xrightarrow[m\to\infty]{} 0$$

by definition of Schauder basis. Then $x = \sum_n c_n e_n$ in $(\mathcal{X}, \|\cdot\|_b)$, and we can deduce that $\{e_n\}$ is a Schauder basis for $(\mathcal{X}, \|\cdot\|_b)$ (Exercise 9.8.4). Also

$$\|P_m\|_b = \sup_{\|x\|_b=1} \|P_m x\|_b = \sup_{\|x\|_b=1} \sup_n \|P_n P_m x\| = \sup_n \sup_{\sup_k \|P_k x\|=1} \|P_n P_m x\|$$
$$= \sup_n \sup_{\sup_k \|P_k x\|=1} \|P_{\min\{m,n\}} x\| \leq 1.$$

So the basis constant for $E$ in $(\mathcal{X}, \|\cdot\|_b)$ is 1.

(iv) We have $\mathrm{b}_E \geq 1$ since $\|P_n\| \geq 1$ for all $n$ as the $P_n$ are projections.

□

The notation $e_n^*$ for the biorhogonal functionals could be confusing at first, since they are not adjoints. But there should be no confusion as the $e_n$ are not operators and so they do not have adjoints.

While we defined the basis projections in terms of the Schauder basis, we can go the other way. That is, given projections with the correct properties we can obtain a Schauder basis. Concretely,

***9.8.3. Lemma.***

*Let $\mathcal{X}$ be a normed space and $\{P_n\}$ a sequence of bounded idempotents. Then $\{P_n\}$ are basis projections for a Schauder basis $\{e_n\}$ if and only if*

*(i)* $\dim P_n\mathcal{X} = n$*;*

*(ii)* $P_n x \to x$ *for all* $x \in \mathcal{X}$*;*

*(iii)* $P_n P_m = P_{\min\{m,n\}}$.

*Proof.* The three properties are automatic given by the definition of Schauder basis. So assume we have a sequence of projections satisfying the three properties. For each $n \in \mathbb{N}$, choose nonzero $e_n \in P_n\mathcal{X} \cap \ker P_{n-1}$, where we write $P_0 = 0$. We then can see, by alternatively applying $P_{n+1}$ and $P_n$, that

$$P_{n+1}x - P_n x \in P_{n+1}\mathcal{X} \cap \ker P_n, \qquad x \in \mathcal{X}.$$

Also, $\dim P_{n+1}\mathcal{X} \cap \ker P_n = 1$, since $\dim P_{n+1}\mathcal{X} = n+1$, $\dim P_n\mathcal{X} = n$, and $P_n\mathcal{X} \subset P_{n+1}\mathcal{X}$; indeed, since an element of $\ker P_n$ is linearly independent with $P_n\mathcal{X}$ (from $P_n$ being an idempotent), we have $P_{n+1}\mathcal{X} = P_n\mathcal{X} \oplus (P_{n+1}\mathcal{X} \cap \ker P_n)$, which gives $n + 1 = n + \dim P_{n+1}\mathcal{X} \cap \ker P_n$. Thus $P_{n+1}x - P_n x =$

$b_{n+1}e_{n+1}$ for some $b_{n+1} \in \mathbb{C}$. Then

$$x = \lim_n P_n x = \lim_n \sum_{k=1}^{n} P_n x - P_{n-1}x = \lim_n \sum_{k=1}^{n} b_n e_n = \sum_{k=1}^{\infty} b_n e_n.$$

So all that remains to see that $\{e_n\}$ is a Schauder basis is to check the uniqueness of the coefficients. Note that $P_n e_k = e + k$ if $k \leq$, and $P_n e_k = 0$ if $k > n$; for $P_n e_k = (P_n P_{n+k})e_n = P_n(P_{n+k}e_n) = 0$. If $x = \sum_{k=1}^{\infty} b'_k e_k$ then, using that each $P_n$ is continuous (being bounded)

$$P_n x = P_n\Big(\sum_{k=1}^{\infty} b'_k e_k\Big) = \sum_{k=1}^{\infty} b'_k P_n e_k = \sum_{k=1}^{n} b'_k e_k.$$

Then $b'_n e_n = P_n x - P_{n-1}x = b_n e_n$ and so the coefficients are unique. □

When $\mathcal{X}$ admits a sequence $\{P_n\} \subset \mathcal{B}(\mathcal{X})$ satisfying (ii) in Lemma 9.8.3 we say that $\mathcal{X}$ has the **Approximation Property**.

*9.8.4. Examples*. We refer to Section 5.1 for the most basic examples of Schauder bases. Here are a few less trivial ones. In $C[0,1]$, the **Faber–Schauder basis** $F$ is constructed as follows. Let $\{t_j\} \subset [0,1]$ be a dense sequence with $t_1 = 0$, and $t_2 = 1$. Define $P_n f$ to be the piecewise linear function $[0,1] \to [0,1]$ such that $(P_n f)(t_j) = f(t_j)$ for $j = 1,\dots,n$. Because $P_{n+1}f$ and $P_n f$ agree on $t_1,\dots,t_n$, we have $P_{n+1}P_n = P_n$. As the range of $P_n$ is $n$-dimensional (this is a routine but interesting exercise if the reader has never done it! See Exercise 9.8.5), all that remains to be seen for the $\{P_n\}$ to be basis projections is—thanks to Lemma 9.8.3—that $P_n x \to x$. This follows rather easily from the fact that each $f \in C[0,1]$ is uniformly continuous (Exercise 1.8.17). So the $\{P_n\}$ determine a Schauder basis for $C[0,1]$ called the Faber–Schauder basis. Since $e_n \in P_n\mathcal{X} \cap \ker P_{n-1}$, it is piecewise linear and it can be normalized so that $e_n(t_n) = 1$ and $e_n(t_j) = 0$ for $j = 1,\dots,n-1$. So $e_1 = 1$, $e_2(t) = t$, and each successive $e_n$ has the shape of a triangle with $e_n(t_n) = 1$ and reaching 0 at the two nearest $t_j$. The construction gives $\|P_n f\| \leq \|f\|$ for all $n$ and all $f$, and so $b_F = 1$.

For $1 \leq p < \infty$ the **Haar basis** is defined by $h_0 = 1$ and

$$h_{m,n} = 1_{[\frac{m-1}{2^n},\frac{m}{2^n}]} - 1_{(\frac{m}{2^n},\frac{m+1}{2^n}]}, \qquad n \in \mathbb{N},\ m = 1,3,5,\dots,2^n-1.$$

In some contexts (and with a different normalization, and sometimes different conventions regarding endpoints) the Haar basis is called the **Haar wavelet**. Another way to define the Haar basis is to notice that

$$h_{m,n}(t) = h_{1,1}(2^n t + 1 - m).$$

These functions are orthogonal in $L^2[0,1]$. Indeed, if $n' \geq n$, then $h_{m,n}h_{m',n'}$ is either 0 or $h_{m',n'}$, and in both cases $\int_0^1 h_{m,n}h_{m',n'} = 0$. In particular, they are linearly independent. Next we claim that the span of $\{h_0\} \cup \{h_{m,n}\}_{m,n}$ is dense in $L^p[0,1]$. First note that

$$1_{[0,\frac{1}{2}]} = \frac{1}{2}\,(h_0 + h_{11}), \qquad\qquad 1_{(\frac{1}{2},1]} = \frac{1}{2}\,(h_0 - h_{11}).$$

Next,

$$1_{[0,1/4]} = \frac{1}{2}\,(1_{[0,\frac{1}{2}]} + h_{12}), \qquad 1_{(\frac{1}{4},\frac{1}{2}]} = \frac{1}{2}\,(1_{[0,\frac{1}{2}]} - h_{12}),$$

and

$$1_{[\frac{1}{2},\frac{3}{4}]} = \frac{1}{2}\,(1_{[\frac{1}{2},\frac{3}{4}]} + h_{22}), \qquad 1_{(\frac{3}{4},1]} = \frac{1}{2}\,(1_{[\frac{1}{2},\frac{3}{4}]} - h_{22}).$$

Going ahead with the same idea, it follows that every $1_L$ with $L$ a dyadic interval is in $\operatorname{span}\{h_0\} \cup \{h_{m,n}\}_{m,n}$. Since any continuous function on $[0,1]$ can be uniformly approximated by linear combinations of characteristics of dyadic intervals (because of the uniform continuity), Proposition 2.8.18 guarantees that

$$\overline{\operatorname{span}}^{\|\cdot\|}\,\{h_0\} \cup \{h_{m,n}\}_{m,n} = L^p[0,1], \qquad\qquad 1 \leq p < \infty.$$

And since the sequence is orthogonal and generating, if we normalize it we get an orthonormal basis of $L^2[0,1]$; this is the most common setting where the term *Haar wavelet* is used. Concretely, $\{h_0\} \cup \{2^{(n-1)/2} h_{m,n}\}_{m,n}$ is an orthonormal basis for $L^2[0,1]$.

It still remains to be seen that the Haar basis is a Schauder basis for $L^p[0,1]$, as to span a dense subset is not enough: for an example of a linearly independent subset of a Banach space that spans a dense subset but is not a Schauder basis, consider the monomials in $C[0,1]$, that span a dense set, but the convergent series on them give the analytic functions, a proper subset.

For simplicity in the rest of the argument let us renumber $\{h_0\} \cup \{h_{m,n}\}_{m,n}$ as $\{h_k\}_{k=0}^\infty$. On $\operatorname{span}\{h_k\}$ we define $P_n\big(\sum_{k=1}^m b_k h_k\big) = \sum_{k=1}^n b_k h_k$, with $b_k = 0$ if $k > m$. Now let $f = \sum_{k=1}^n b_k h_k$, $g = \sum_{k=1}^{n+1} b_k h_k$. Then either $g = f$, or $g - f = b_{n+1} h_{n+1}$. This is a dyadic interval $L$ where $f$ is constant, say it takes the value $b$, and $g$ takes two values, say $b+a$ and $b-a$. When $1 \leq p < \infty$ the function $t \longmapsto |t|^p$ is convex. So

$$|b|^p = \left|\frac{1}{2}(b+a) + \frac{1}{2}(b-a)\right|^p \leq \frac{1}{2}\,|b+a|^p + \frac{1}{2}\,|b-a|^p.$$

Then, writing $L = L_1 \cup L_2$ as the two halves of the interval determined by $h_{n+1}$,

$$\begin{aligned}
\|f\|_p^p &= \int_L |b|^p + \int_{[0,1]\setminus L} |f|^p = |b|^p\, m(L) + \int_{[0,1]\setminus L} |f|^p \\
&\leq \frac{1}{2}\,|b+a|^p\, m(L) + \frac{1}{2}\,|b-a|^p\, m(L) + \int_{[0,1]\setminus L} |g|^p \\
&= |b+a|^p\, m(L_1) + |b-a|^p\, m(L_2) + \int_{[0,1]\setminus L} |g|^p \\
&= \|g\|_p^p.
\end{aligned}$$

Thus $\|P_n g\|_p = \|f\|_p \leq \|g\|_p$. Inductively, this shows that $\|P_n g\|_p \leq \|g\|_p$ for all $g \in \operatorname{span}\{h_k\}$. So $\|P_n\| = 1$ and it extends to $\overline{\operatorname{span}}^{\|\cdot\|}\,\{h_k\}$. Therefore the

$\{P_n\}$ are uniformly bounded projections that trivially satisfy $\dim P_n\mathcal{X} = n$ and $P_mP_n = P_{\min\{m,n\}}$; and, being bounded, they extend to all of $L^p[0,1]$. By Lemma 9.8.3, all that remains to see is that $\lim_n P_nx = x$. By the density of $\operatorname{span}\{h_k\}$, for any $\varepsilon > 0$ there exists $m$ and coefficients such that if $x_n = \sum_{k=1}^m b_kh_k$ then $\|x - x_n\| < \varepsilon/2$. Then

$$\|x-P_nx\| \leq \|x-x_n\|+\|x_n-P_nx\| = \|x-x_n\|+\|P_n(x_n-x)\| \leq 2\|x-x_n\| < \varepsilon.$$

So $P_nx \to x$ and this concludes the argument that the $\{h_k\}$ are a Schauder basis.

We mentioned on page 608 that the presence of a Schauder basis gives the equality $\overline{\mathcal{F}(\mathcal{X},\mathcal{Y})} = \mathcal{K}(\mathcal{X},\mathcal{Y})$. Indeed,

***9.8.5. Proposition.***

*Let $\mathcal{X},\mathcal{Y}$ be Banach spaces such that $\mathcal{X}$ has a Schauder basis. Then $\overline{\mathcal{F}(\mathcal{X},\mathcal{Y})} = \mathcal{K}(\mathcal{X},\mathcal{Y})$.*

*Proof.* Lemma 9.8.3 tells us that if $\{P_n\}$ are the basis projections associated with the Schauder basis, then $P_nx \to x$ for all $x \in \mathcal{X}$. Given $T \in \mathcal{K}(\mathcal{X},\mathcal{Y})$, we will show that $P_nT \to T$. We already have that this occurs pointwise, but we need the limit to be uniform. Fix $\varepsilon > 0$. Because $T$ is compact, $T\overline{B_1^{\mathcal{X}}(0)}$ admits an $\varepsilon$ net; so there exist $x_1,\dots,x_m$ such that

$$T\overline{B_1^{\mathcal{X}}(0)} \subset \bigcup_{j=1}^m B_\varepsilon(Tx_j).$$

Now, given $x \in \overline{B_1^{\mathcal{X}}(0)}$ there exists $j$ such that $\|Tx - Tx_j\| < \varepsilon$. Then, with $b_E = \sup\{\|P_n\| : n\}$,

$$\begin{aligned}\|Tx - P_nTx\| &\leq \|Tx - Tx_j\| + \|Tx_j - P_nTx_j\| + \|P_nTx_j - P_nTx\| \\ &\leq (1+b_E)\varepsilon/3 + \|Tx_j - P_nTx_j\|.\end{aligned}$$

By the approximation property, there exists $n_0$ such that $\|Tx_j - P_nTx_j\| < \varepsilon$ for all $n \geq n_0$. Therefore, for all $n \geq n_0$ we have $\|Tx - P_nTx\| \leq (2+b_E)\varepsilon$. And this was for all $x$ in the unit ball, so $\|T - P_nT\| \leq (2+b_E)\varepsilon$, showing that $P_nT \to T$. $\square$

**9.8.2. Basic Sequences.** We now turn our attention to basic sequences.

**9.8.6. Proposition.**

*Let $\mathcal{X}$ be a Banach space and $E = \{e_n\} \subset \mathcal{X}$ with $e_n \neq 0$ for all $n$. The following statements are equivalent:*

*(i) $\{e_n\}$ is a basic sequence;*

*(ii) there exists $c > 0$ such that for all $m, n \in \mathbb{N}$ with $m \leq n$ and for all $\{c_n\} \subset \mathbb{C}$,*

$$\Big\|\sum_{j=1}^{m} c_j e_j\Big\| \leq c \Big\|\sum_{j=1}^{n} c_j e_j\Big\|. \tag{9.39}$$

*Proof.* (i) $\implies$ (ii) Write $\{P_n\}$ for the basic projections. Then

$$\Big\|\sum_{j=1}^{m} c_j e_j\Big\| = \Big\|P_m\Big(\sum_{j=1}^{n} c_j e_j\Big)\Big\| \leq \mathrm{b}_E \Big\|\sum_{j=1}^{n} c_j e_j\Big\|.$$

(ii) $\implies$ (i) Let $\mathcal{X}_0 = \operatorname{span}\{e_n\}$. The inequality (9.39) guarantees that the $\{e_n\}$ are linearly independent (Exercise 9.8.6). Define $P_m : \mathcal{X}_0 \to \operatorname{span}\{e_1, \dots, e_m\}$ by

$$P_m\Big(\sum_{j=1}^{n} a_j e_j\Big) = \sum_{j=1}^{\min\{m,n\}} a_j e_j, \qquad m, n \in \mathbb{N}.$$

From (9.39) we have that $\|P_m\| \leq \max\{c, 1\}$ for all $m$; then it extends to a unique $P_m \in \mathcal{B}(\overline{\mathcal{X}_0})$ by Proposition 6.1.9. By Lemma 9.8.3 we have that $\{e_n\}$ is a Schauder basis for $\overline{\mathcal{X}_0}$. ☐

**9.8.7. Proposition.**

*Let $\mathcal{X}$ be a Banach space and $\{e_n\}$ a basic sequence. Let $\{f_n\} \subset \mathcal{X}$ with $\sum_n \|e_n - f_n\| \, \|e_n^*\| < 1$. Then*

*(i) the sequence $\{f_n\}$ is basic and equivalent to $\{e_n\}$;*

*(ii) $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$ is complemented in $\mathcal{X}$ if and only if $\overline{\operatorname{span}}^{\|\cdot\|}\{f_n\}$ is;*

*(iii) $\{e_n\}$ is a Schauder basis of $\mathcal{X}$ if and only if $\{f_n\}$ is. In such case, $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n^*\} = \overline{\operatorname{span}}^{\|\cdot\|}\{f_n^*\}$.*

*Proof.* (i) Write $c = \sum_n \|e_n - f_n\| \, \|e_n^*\|$; then $c < 1$ by hypothesis. By Proposition 9.8.2 each $e_n^*$ is bounded on $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$; using Hahn–Banach (Corollary 5.7.6) we can then extend each $e_n^*$ from $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$ to all of $\mathcal{X}$

while keeping their norm. Define a linear operator $S : \mathcal{X} \to \mathcal{X}$ by

$$Sx = \sum_{n=1}^{\infty} e_n^*(x)\,(e_n - f_n).$$

We have

$$\|Sx\| \le \sum_{n=1}^{\infty} |e_n^*(x)|\,\|e_n - f_n\| \le \sum_{n=1}^{\infty} \|e_n\|\,\|e_n - f_n\|\,\|x\| \le c\|x\|,$$

so $S$ is bounded with $\|S\| \le c < 1$. With $T = I_{\mathcal{X}} - S$, we have $\|x - Tx\| = \|Sx\| \le c\|x\|$ and, via the reverse triangle inequality, $\|Tx\| \ge (1 - c)\|x\|$. Therefore $T$ is bounded below, thus injective and with closed range (Exercise 9.4.1). Suppose that $T\mathcal{X} \subsetneq \mathcal{X}$. Being a proper closed subspace, by Riesz's Lemma (Lemma 5.2.8) there exists $x \in \mathcal{X}$ with $\|x\| = 1$ and $\operatorname{dist}(x, T\mathcal{X}) > c$. In particular $\|Sx\| = \|x - Tx\| > c$, a contradiction. Thus $T$ is surjective, and so an automorphism of $\mathcal{X}$. Since $Te_n = f_n$ for all $n$, this establishes (i).

(ii) Since $T$ is an automorphism (that is, bijective, bounded, with bounded inverse) of $\mathcal{X}$, it will map complemented subspaces to complemented subspaces. Explicitly, if $P$ is a bounded projection onto a subspace $\mathcal{X}_0$, then $TPT^{-1}$ is a bounded projection onto $T\mathcal{X}_0$. The argument is complete since $T$ maps $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$ onto $\overline{\operatorname{span}}^{\|\cdot\|}\{f_n\}$.

(iii) As $T$ is an automorphism that maps $\{e_n\}$ to $\{f_n\}$, one will be a Schauder basis if and only if the other is. Fix $h \in \mathbb{N}$ and let $g_k = \sum_{j=1}^{k} f_h^*(e_j)e_j^* \in \operatorname{span}\{e_n^*\}$. Let $x \in \mathcal{X}$ with $\|x\| \le 1$. Then, since $f_h^*$ is bounded,

$$f_h^*(x) = f_h^*\Big(\sum_j c_j e_j\Big) = \lim_k \sum_{j=1}^{k} c_j\, f_h^*(e_j) = \lim_k \sum_{j=1}^{k} f_h^*(e_j)\, e_j^*(x) = \lim_k g_k(x).$$

Now, if $k \ge h$,

$$|(f_h j^* - g_k)x| = \Big|\sum_{j=k+1}^{\infty} f_h^*(e_j)\, e_j^*(x)\Big| = \Big|\sum_{j=k+1}^{\infty} f_h^*(e_j - f_j)\, e_j^*(x)\Big|$$

$$\le \|f_h^*\| \sum_{j=k+1}^{\infty} \|e_j - f_j\|\,\|e_j^*\| \xrightarrow[k\to\infty]{} 0.$$

This shows that $f_h^* \in \overline{\operatorname{span}}^{\|\cdot\|}\{e_n^*\}$. As the roles of the $e_n^*$ and $f_n^*$ can be exchanged, we get that $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n^*\} = \overline{\operatorname{span}}^{\|\cdot\|}\{f_n^*\}$. □

**9.8.8. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. Recall that two basic sequences $E = \{e_n\} \subset \mathcal{X}$ and $F = \{f_n\} \subset \mathcal{Y}$ are **equivalent**, denoted $E \sim F$, if for any $\{a_n\} \subset \mathbb{C}$ we have that $\sum_n a_n e_n$ converges if and only if $\sum_n a_n f_n$ converges. The* *sequences $E$ and $F$ are* ***isometrically equivalent*** *if for any $a_1, \ldots, a_m \in \mathbb{C}$ we have*

$$\Big\| \sum_{n=1}^{m} a_n e_n \Big\| = \Big\| \sum_{n=1}^{m} a_n f_n \Big\|.$$

*A sequence $\{v_n\} \subset \mathcal{X}$ is a* ***block basic sequence*** *for $E$ if there exist integers $0 = m_0 < m_1 < \cdots$ and $\{a_n\} \subset \mathbb{C}$ with*

$$v_n = \sum_{j=m_{n-1}+1}^{m_n} a_j e_j.$$

*A basic sequence $E \subset \mathcal{X}$ is* ***complemented*** *if $\overline{\operatorname{span}}^{\|\cdot\|} E$ is complemented in $\mathcal{X}$.*

At first sight it is hard to imagine how a block basic sequence can be useful, but we will soon see.

***9.8.9. Proposition.***

*Let $E = \{e_n\} \subset \mathcal{X}$ and $F = \{f_n\} \subset \mathcal{Y}$ be basic sequences of the Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ respectively. The following statements are equivalent:*

*(i)* $E \sim F$;

*(ii)* *the map $S : e_n \longmapsto f_n$ extends to a bounded linear isomorphism between the spaces $\overline{\operatorname{span}}^{\|\cdot\|} \{e_n\}$ and $\overline{\operatorname{span}}^{\|\cdot\|} \{f_n\}$;*

*(iii)* *there exists $c > 0$ such that for all $a_1, \ldots, a_m \in \mathbb{C}$*

$$\frac{1}{c} \Big\| \sum_{n=1}^{m} a_n f_n \Big\| \leq \Big\| \sum_{n=1}^{m} a_n e_n \Big\| \leq c \Big\| \sum_{n=1}^{m} a_n f_n \Big\|. \tag{9.40}$$

*Proof.* (i) $\Longrightarrow$ (ii) To simplify notation we may assume, without loss of generality, that $\mathcal{X} = \overline{\operatorname{span}}^{\|\cdot\|} \{e_n\}$ and $\mathcal{Y} = \overline{\operatorname{span}}^{\|\cdot\|} \{f_n\}$. Let $S : \mathcal{X} \to \mathcal{Y}$ be given by $S\big(\sum_{j=1}^{\infty} a_j e_j\big) = \sum_{j=1}^{\infty} a_j f_j$. Then $S$ is bijective (Exercise 9.8.9). To show that $S$ is bounded we use the Closed Graph Theorem (6.3.12). Indeed, suppose that $\{x_n\} \subset \mathcal{X}$ converges to $x \in \mathcal{X}$ and $Sx_n \to y \in \mathcal{Y}$. We have

$$x_n = \sum_{j=1}^{\infty} e_j^*(x_n)\, e_j, \qquad x = \sum_{j=1}^{\infty} e_j^*(x)\, e_j.$$

Since the biorthogonal functionals are bounded (Proposition 9.8.2), $e_j^*(x_n) \to e_j^*(x)$ and $f_j^*(Sx_n) \to f_j^*(y)$ for all $j$. By definition of $S$, $f_j^*(Sz) = e_j^*(z)$ for all $z \in \mathcal{X}$. Thus

$$f_j^*(Sx) = e_j^*(x) = \lim_n e_j^*(x_n) = \lim_n f_j^*(Sx_n) = f_j^*(y)$$

for all $j$; since the sequences are basic we are guaranteed unique coefficients, and so $Sx = y$ and $S$ is bounded. Its inverse is then bounded by the Inverse Mapping Theorem (6.3.6).

(ii) $\Longrightarrow$ (iii) Since $S$ and $S^{-1}$ are bounded,

$$\left\| \sum_{n=1}^{\infty} a_n f_n \right\| = \left\| S\Big( \sum_{n=1}^{\infty} a_n e_n \Big) \right\| \le \|S\| \left\| \sum_{n=1}^{\infty} a_n e_n \right\|$$

and

$$\left\| \sum_{n=1}^{\infty} a_n e_n \right\| = \left\| S^{-1}\Big( \sum_{n=1}^{\infty} a_n f_n \Big) \right\| \le \|S^{-1}\| \left\| \sum_{n=1}^{\infty} a_n f_n \right\|.$$

Then we can take $c = \max\{\|S\|, \|S^{-1}\|\}$.

(iii) $\Longrightarrow$ (i) The inequalities (9.40) guarantee that the tail of one series will be small if and only if the tail of the other series is small. $\square$

In tune with their name, block basic sequences are actually basic sequences.

**9.8.10. Lemma.**

*Let $\mathcal{X}$ be a Banach space with Schauder basis $E = \{e_n\}$ and basis constant $\mathrm{b}_E$. Let $V = \{v_k\}$ be a block sequence for $E$. Then $V$ is a basic sequence with $\mathrm{b}_V \le \mathrm{b}_E$.*

*Proof.* If $v_k = \sum_{j=m_{k-1}+1}^{m_k} a_j e_j$ and $\{c_k\} \subset \mathbb{C}$, $n_1, n_2 \in \mathbb{N}$ with $n_1 \le n_2$, we define

$$d_j = c_k a_j, \qquad m_{k-1} + 1 \le j \le m_k.$$

This gives

$$\begin{aligned}\left\| \sum_{k=1}^{n_1} c_k v_k \right\| &= \left\| \sum_{k=1}^{n_1} \sum_{j=m_{k-1}+1}^{m_k} c_k a_j e_j \right\| = \left\| \sum_{j=1}^{m_{n_1}} d_j e_j \right\| \\ &\le \mathrm{b}_E \left\| \sum_{j=1}^{m_{n_2}} d_j e_j \right\| = \mathrm{b}_E \left\| \sum_{k=1}^{n_2} c_k v_k \right\|.\end{aligned}$$

Then Proposition 9.8.6 gives us that $V$ is a basic sequence. $\square$

**9.8.11. Proposition.**

*Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a basic sequence. Let $Y = \{y_n\} \subset \mathcal{X}$. If*

$$\theta = 2\,\mathrm{b}_E \sum_{k=1}^{\infty} \frac{\|e_n - y_n\|}{\|e_n\|} < 1,$$

*then* $\{y_n\}$ *is a basic sequence, and*

(i) $\{e_n\}$ *and* $\{y_n\}$ *are equivalent;*

(ii) $\mathrm{b}_Y \leq \mathrm{b}_E(1+\theta)(1-\theta)^{-1}$;

(iii) *if* $\{e_n\}$ *is complemented, so is* $\{y_n\}$.

*Proof.* We know from Proposition 9.8.2 that the biorthogonal functionals $\{e_n^*\}$ are bounded with $\|e_n^*\| \leq 2\,\mathrm{b}_E\,\|e_n\|^{-1}$. Via Hahn–Banach (Corollary 5.7.6) we can assume that $e_n^* \in \mathcal{X}^*$. Define $S : \mathcal{X} \to \mathcal{X}$ by

$$Sx = x + \sum_{k=1}^{\infty} e_n^*(x)\,(y_n - e_n).$$

Then $Se_n = y_n$ and

$$\|S\| \leq 1 + 2b_E \sum_{k=1}^{\infty} \frac{\|e_n - y_n\|}{\|e_n\|} = 1 + \theta.$$

So $S \in \mathcal{B}(\mathcal{X})$. Also,

$$\|x - Sx\| \leq \sum_{k=1}^{\infty} \|x\|\,\|e_n^*\|\,\|y_n - e_n\| \leq \theta\,\|x\|.$$

Then $\|I - S\| \leq \theta < 1$ and $S$ is invertible by Proposition 6.2.3, with

$$\|S^{-1}\| = \left\| \sum_{k=0}^{\infty} (I-S)^k \right\| \leq \sum_{k=0}^{\infty} \theta^k = (1-\theta)^{-1}.$$

Since $S$ is an isomorphism between $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$ and $\overline{\operatorname{span}}^{\|\cdot\|}\{y_n\}$ with $Se_n = y_n$, we get that $\{y_n\}$ is a basic sequence. We get that both sequences are equivalent by Proposition 9.8.9 and the basic projections for $\{y_n\}$ are $Q_m = SP_mS^{-1}$. Then

$$\|Q_m\| \leq \|P_m\|\,\|S\|\,\|S^{-1}\| \leq \mathrm{b}_E\,(1+\theta)(1-\theta)^{-1}.$$

When $\{e_n\}$ is complemented and $P$ is a bounded projection onto $\overline{\operatorname{span}}^{\|\cdot\|}\{e_n\}$, we have that $Q = SPS^{-1}$ is a bounded projection onto $\overline{\operatorname{span}}^{\|\cdot\|}\{y_n\}$ and hence $\{y_n\}$ is complemented. $\square$

The next proposition is fairly technical, and it is hard to see why one would bother. But its consequences are far reaching.

**9.8.12. Theorem.** *(The Bessaga–Pełczyński Selection Principle)*

*Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis for $\mathcal{X}$. Let $\{x_n\} \subset \mathcal{X}$ such that*

*(i) there exists $c > 0$ with $\|x_n\| \geq c$ for all $n$;*

*(ii) $\lim_n e_k^*(x_n) = 0$ for all $k$.*

*Then there exists a subsequence $E' = \{x_{n_k}\}$, which is a basic sequence equivalent to a block basic sequence $Y = \{y_n\}$ of $\{e_n\}$. If $Y$ is complemented, then so is $E'$. If $\varepsilon > 0$ is fixed, the subsequence can be chosen so that $\mathrm{b}_{E'} \leq \mathrm{b}_E + \varepsilon$.*

*Proof.* As usual we denote by $P_n$ the basis projections associated with $E$. Fix $\delta$ with $0 < \delta < \frac{1}{4}$. Let $n_1 = 1$, $r_0 = 0$. Since $\lim_r P_r x = x$ for all $x$, there exists $r_1 \in \mathbb{N}$ such that

$$\|x_{n_1} - P_{r_1} x_{n_1}\| < \frac{\delta c}{2\,\mathrm{b}_E}.$$

As $P_{r_1} x_n$ is a finite linear combination of $e_1, \ldots, e_{r_1}$, the hypothesis (ii) implies that $P_{r_1} x_n \underset{n}{\to} 0$. So there exists $n_2 > n_1$ such that

$$\|P_{r_1} x_{n_2}\| < \frac{\delta^2 c}{2\,\mathrm{b}_E}.$$

Now we can choose $r_2 > r_1$ such that

$$\|x_{n_2} - P_{r_2} x_{n_2}\| < \frac{\delta^2 c}{2\,\mathrm{b}_E}.$$

Continuing inductively, we get a subsequence $\{x_{n_k}\} \subset \mathcal{X}$ and an increasing sequence of integers $\{r_k\}$, with $r_0 = 0$, and with

$$\|x_{n_k} - P_{r_k} x_{n_k}\| < \frac{\delta^k c}{2\,\mathrm{b}_E}, \qquad \|P_{r_{k-1}} x_{n_k}\| < \frac{\delta^k c}{2\,\mathrm{b}_E}.$$

Now we define $y_k = P_{r_k} x_{n_k} - P_{r_{k-1}} x_{n_k}$. Then $Y = \{y_k\}$ is a block basic sequence (each $y_k$ is a linear combination of $e_1, \ldots, e_{r_k}$); and by Lemma 9.8.10 it is a basic sequence with $\mathrm{b}_Y \leq \mathrm{b}_E$. We have

$$\|y_k - x_{n_k}\| \leq \|x_{n_k} - P_{r_k} x_{n_k}\| + \|P_{r_{k-1}} x_{n_k}\| \leq \frac{\delta^k c}{2\,\mathrm{b}_E} + \frac{\delta^k c}{2\,\mathrm{b}_E} = \frac{\delta^k c}{\mathrm{b}_E}.$$

Thus

$$\|y_k\| \geq \|x_{n_k}\| - \frac{\delta^k c}{\mathrm{b}_E} \geq c - \frac{\delta c}{\mathrm{b}_E} \geq c(1-\delta),$$

and

$$2\,\mathrm{b}_E \sum_{k=1}^{\infty} \frac{\|y_k - x_{n_k}\|}{\|y_k\|} \leq 2\,\mathrm{b}_E \; \frac{1}{c(1-\delta)} \sum_{k=1}^{\infty} \frac{\delta^k c}{\mathrm{b}_E}$$

$$= \frac{2}{1-\delta} \sum_{k=1}^{\infty} \delta^k = \frac{2\delta}{(1-\delta)^2} < \frac{8}{9} < 1.$$

Proposition 9.8.11 then guarantees that $\{x_{n_k}\}$ is a basic sequence equivalent to $Y$ (and complemented if $Y$ is), and

$$\mathrm{b}_{E'} \leq \mathrm{b}_E (1 + 2\delta(1-\delta)^{-2})(1 - 2\delta(1-\delta)^{-2})^{-1}.$$

As $\delta$ can be chosen arbitrarily close to 0, the expression above can be made arbitrarily close to $\mathrm{b}_E$. $\square$

An immediate consequence of Theorem 9.8.12 is that any separable Banach space contains a basic sequence. This is not as trivial as it seems, since there exist separable Banach spaces without a Schauder basis.

***9.8.13. Corollary.***

*Let $\mathcal{X}$ be an infinite-dimensional separable Banach space, and $\varepsilon > 0$. Then there exists $E = \{e_n\} \subset \mathcal{X}$ such that $E$ is a basic sequence and $b_E \leq 1 + \varepsilon$.*

*Proof.* By Banach–Mazur (Theorem 7.6.6) we may assume that $\mathcal{X} \subset C[0,1]$. Now $\mathcal{X}$ lives in a space with a Schauder basis, say $G = \{g_n\}$ (we know for sure there is a Schauder basis, the Faber–Schauder basis from Examples 9.8.4, with $b_G = 1$). Let

$$\mathcal{X}_n = \mathcal{X} \cap \bigcap_{j=1}^{n} \ker g_n^*.$$

By construction we have $\mathcal{X}_1 \supset \mathcal{X}_2 \supset \cdots$, and $\mathcal{X}_n \neq \{0\}$ for all $n$ by Proposition 5.5.12 since $\dim \mathcal{X} = \infty$. By passing to a subsequence if needed we may assume that all inclusions are proper. Indeed, if that were not the case there would only be finitely many proper inclusions, and so there would exist $n$ such that $\mathcal{X}_m = \mathcal{X}_n$ for all $m \geq n$. Then

$$x - \sum_{j=1}^{n} g_j^*(x)\, g_j \in \ker g_m^*$$

for all $m \geq n$, implying that $\dim \mathcal{X} < \infty$, a contradiction. After relabelling the subsequence as $\{\mathcal{X}_n\}$, we choose $f_n \in \mathcal{X}_n \setminus \mathcal{X}_{n+1}$ with $\|f_n\| = 1$ for all $k$. By construction, $g_m^*(f_n) = 0$ if $n \geq m$. Thus $\lim_n g_m^*(f_n) = 0$; together with $\|f_n\| = 1$ for all $k$, Theorem 9.8.12 gives us a subsequence $E = \{e_n\}$ that is a basic sequence with $b_E \leq b_F + \varepsilon = 1 + \varepsilon$. $\square$

We now have tools to focus on our old friends $\ell^p(\mathbb{N})$, $1 \leq p < \infty$, and $c_0$. Recall that for these spaces the canonical basis $\{e_n\}$ is a normalized Schauder basis. The following lemma is rather elementary even if its proof is long. It will open the door to nice results, though.

**9.8.14. Lemma.**

*Let $\mathcal{X}$ be $c_0$ or $\ell^p(\mathbb{N})$, $1 \le p < \infty$. Let $V = \{v_n\}$ be a normalized block basic sequence. Then $V$ is isometrically equivalent to the canonical basis, and $\overline{\operatorname{span}}^{\|\cdot\|} V$ is complemented, via a contractive projection.*

*Proof.* Consider first the case where $\mathcal{X} = \ell^p(\mathbb{N})$. By assumption,

$$v_k = \sum_{j=m_{k-1}+1}^{m_k} c_j e_j, \qquad k \in \mathbb{N},$$

where $m_0 = 0$, $m_k \in \mathbb{N}$, and $m_k < m_{k+1}$ for all $k$. We are also assuming that

$$1 = \|v_k\|_p^p = \sum_{j=m_{k-1}+1}^{m_k} |c_j|^p.$$

Then for $m \in \mathbb{N}$ and $b_1, \ldots, b_m \in \mathbb{C}$

$$\begin{aligned} \left\| \sum_{k=1}^m b_k v_k \right\|_p^p &= \left\| \sum_{k=1}^m \sum_{j=m_{k-1}+1}^{m_k} b_k c_j e_j \right\|_p^p = \sum_{k=1}^m \sum_{j=m_{k-1}+1}^{m_k} |b_k|^p \, |c_j|^p \\ &= \sum_{k=1}^m |b_k|^p = \left\| \sum_{k=1}^m b_k e_k \right\|_p^p . \end{aligned}$$

So the linear map $S : \overline{\operatorname{span}}^{\|\cdot\|} V \to \mathcal{X}$ induced by $v_k \longmapsto e_k$ is isometric with dense range and hence invertible (see Exercise 9.8.7). In the case of $c_0$, the assumption is

$$1 = \|v_k\|_\infty = \max\{|c_j| : \; m_{k-1} + 1 \le j \le m_k\}, \qquad k \in \mathbb{N},$$

and the computation becomes

$$\begin{aligned} \left\| \sum_{k=1}^m b_k v_k \right\|_\infty &= \left\| \sum_{k=1}^m \sum_{j=m_{k-1}+1}^{m_k} b_k c_j e_j \right\|_\infty \\ &= \max\{|b_k| \, |c_j| : \; 1 \le k \le m, \; m_{k-1} + 1 \le j \le m_k\} \\ &= \max\{|b_k| : \; 1 \le k \le m\} = \left\| \sum_{k=1}^m b_k e_k \right\|_\infty \end{aligned}$$

and again the linear map $S$ induced by $v_k \longmapsto e_k$ is isometric and invertible, giving us the isometric equivalence between $\{v_k\}$ and $\{e_k\}$.

The last thing we need to prove is that $\overline{\operatorname{span}}^{\|\cdot\|} V$ is complementable via a contractive projection. First we define linear functionals $v_k^* \in \mathcal{X}^*$ by

$$v_k^* = \sum_{j=m_{k-1}+1}^{m_k} b_j e_j^*, \qquad k \in \mathbb{N},$$

where the scalars $b_j$ are chosen in such a way that

$$\sum_{j=m_{k-1}+1}^{m_k} b_j c_j = 1, \qquad \Big\| \sum_{j=m_{k-1}+1}^{m_k} b_j e_j^* \Big\|_{\mathcal{X}^*} = 1.$$

The existence of such scalars follows from Proposition 5.6.1. The choice of the coefficients $b_j$ guarantees that $\{v_k^*\}$ are biorthogonal to $\{v_k\}$ and that $\|v_k^*\| = 1$ for all $k$.

We define $P : c_{00} \to \overline{\operatorname{span}}^{\|\cdot\|} V$ by (note that the sum is finite)

$$Px = \sum_{k=1}^{\infty} v_k^*(x)\, v_k.$$

The biorthogonality of $\{v_k^*\}$ gives $P^2 = P$. If we show that $\|Px\| \le \|x\|$, then—as its domain is dense—$P$ will extend as a contraction to all of $\mathcal{X}$, that will still be a projection. Since bounded projections have closed range (Exercise 5.3.5) and the range of $P$ contains $\operatorname{span} V$, we will get that $P\mathcal{X} = \overline{\operatorname{span}}^{\|\cdot\|} V$. To show that $P$ is a contraction, let $x \in c_{00}$. We have, in the $\ell^p(\mathbb{N})$ case, (notice that all sums below are finite so we do not have to discuss convergence)

$$\begin{aligned} \|Px\|_p &= \Big\| \sum_{k=1}^{\infty} v_k^*(x)\, v_k \Big\|_p = \Big\| \sum_{k=1}^{\infty} v_k^*(x)\, e_k \Big\|_p = \bigg( \sum_{k=1}^{\infty} \Big| \sum_{j=m_{k-1}+1}^{m_k} b_j e_j^*(x) \Big|^p \bigg)^{1/p} \\ &\le \bigg( \sum_{k=1}^{\infty} \Big\| \sum_{j=m_{k-1}+1}^{m_k} b_j e_j^* \Big\|_{\mathcal{X}^*}^p \Big( \sum_{j=m_{k-1}+1}^{m_k} |x(j)|^p \Big) \bigg)^{1/p} \\ &= \bigg( \sum_{k=1}^{\infty} \sum_{j=m_{k-1}+1}^{m_k} |x(j)|^p \bigg)^{1/p} = \|x\|_p. \end{aligned}$$

When $\mathcal{X} = c_0$ we have $\sum_{j=m_{k-1}+1}^{m_k} |b_j| = 1$ (since $\mathcal{X}^* = \ell^1(\mathbb{N})$) for all $k$, and then

$$\begin{aligned} \|Px\|_\infty &= \Big\| \sum_{k=1}^{\infty} v_k^*(x)\, v_k \Big\|_\infty = \Big\| \sum_{k=1}^{\infty} v_k^*(x)\, e_k \Big\|_\infty = \max_k \Big| \sum_{j=m_{k-1}+1}^{m_k} b_j e_j^*(x) \Big| \\ &\le \max_k \sum_{j=m_{k-1}+1}^{m_k} |b_j| \max_{m_{k-1}+1 \le j \le m_k} x(j) = \max_j |x(j)| = \|x\|_\infty. \qquad \square \end{aligned}$$

Let us put Lemma 9.8.14 to use.

**9.8.15. Proposition.**

*Let $\mathcal{X}$ be either $\ell^p(\mathbb{N})$ for $1 \leq p < \infty$, or $c_0$. Let $\{x_n\} \subset \mathcal{X}$ be a normalized sequence such that $\lim_n x_n(j) = 0$ for all $j$. Then there exists a subsequence $\{x_{n_k}\}$ that is a basic sequence, equivalent to the canonical basis, and such that $\overline{\operatorname{span}}^{\|\cdot\|}\{x_{n_k}\}$ is complemented.*

*Proof.* By Theorem 9.8.12 there exists a subsequence $E' = \{x_{n_k}\}$ and a block basic sequence $Y = \{y_n\}$ of $\{e_n\}$ such that $E'$ and $Y$ are equivalent, and such that $E'$ will be complemented if $Y$ is; because $\{x_{n_k}\}$ is normalized and $\{y_k\}$ equivalent to it, there exists $c > 0$ with $c^{-1} \leq \|y_k\| \leq c$ for all $k$. Let $S$ be an invertible map with $Sx_{n_k} = y_k$. We have, since $\|x_{n_k}\| = 1$,

$$\|S^{-1}\|^{-1} \leq \|y_k\| \leq \|S\|. \tag{9.41}$$

Applying Lemma 9.8.14 to the normalized block basic sequence $\{y_k/\|y_k\|\}$ we get that $\{y_k/\|y_k\|\}$ is complemented, and so $\{y_k\}$ is complemented since they span the same space. Hence $\{x_{n_k}\}$ is complemented. Lemma 9.8.14 also gives us a linear isometry $W : \mathcal{X} \to \overline{\operatorname{span}}^{\|\cdot\|} Y$ with $We_k = y_k/\|y_k\|$. Tweaking this isometry with the coefficients $\{\|y_k\|\}$ we get a bounded bijection $\mathcal{X} \to \overline{\operatorname{span}}^{\|\cdot\|} Y$ (likely not isometric) that maps $e_k$ to $y_k$. So $Y$ is equivalent to the canonical basis, and thus $\{x_{n_k}\}$ is equivalent to the canonical basis. $\square$

While Proposition 9.8.15 might look a bit obscure for the amount of work we invested in obtaining it, it is the main ingredient in the proof of Pitt's Theorem below.

**9.8.16. Theorem. (Pitt)**

*Let $\mathcal{X}$ be either $\ell^p(\mathbb{N})$ for $1 \leq p < \infty$, or $c_0$. Let $\mathcal{X}_0 \subset \mathcal{X}$ be a closed subspace, and $T \in \mathcal{B}(\mathcal{X}_0, \ell^{p'}(\mathbb{N}))$ with $1 \leq p' < \infty$ and $p' < p$ if $\mathcal{X} = \ell^p(\mathbb{N})$. Then $T$ is compact.*

*Proof.* Consider first the case where $\mathcal{X} = \ell^p(\mathbb{N})$. Since $p > p' \geq 1$ and so $\mathcal{X}$ is reflexive, by Proposition 9.6.5 it is enough to show that $T$ maps weakly convergent bounded sequences to norm convergent sequences (we can require bounded for weakly convergent sequences are bounded, Proposition 7.1.11).

Let $\{x_n\}$ such that $x_n \xrightarrow{\text{weak}} 0$, and suppose that $Tx_n$ does not converge to 0. By passing to a subsequence if needed we may assume that $\|Tx_n\| \geq c > 0$ for some $c$. It follows that $x_n$ does not converge to zero in norm, so we may choose a further subsequence (that we still denote by $\{x_n\}$) such that $\|x_n\| \geq$

$\delta > 0$ for some $\delta > 0$. Then we can replace each $x_n$ with $x_n/\|x_n\|$ and still have that $x_n \xrightarrow{\text{weak}} 0$ and $\|Tx_n\| \geq c'$ for some $c' > 0$, and now with $\|x_n\| = 1$. By Proposition 9.8.15 there is a further subsequence—that we still denote by $\{x_n\}$—that is a complemented basic sequence equivalent to the canonical basis in $\ell^p(\mathbb{N})$. We also have that $Tx_n \xrightarrow{\text{weak}} 0$, since $f \circ T \in \mathcal{X}_0^*$ for all $f \in \ell^{p'}(\mathbb{N})^*$. We can then apply Proposition 9.8.15 to $\{Tx_n/\|Tx_n\|\}$ to extract yet one more subsequence—that we keep denoting without new subscripts—such that $\{Tx_n/\|Tx_n\|\}$ is a complemented basic sequence equivalent to the canonical basis in $\ell^{p'}(\mathbb{N})$. Because $\|Tx_n\| \geq c$ for all $n$, $\{Tx_n\}$ is also a complemented basic sequence equivalent to the canonical basis in $\ell^{p'}(\mathbb{N})$. We have thus shown, via Proposition 9.8.9, that there exist bounded invertible linear maps

$$\ell^p(\mathbb{N}) \xrightarrow{R} \overline{\operatorname{span}}^{\|\cdot\|}\{x_n\} \xrightarrow{T} \overline{\operatorname{span}}^{\|\cdot\|}\{Tx_n\} \xrightarrow{S} \ell^{p'}(\mathbb{N})$$

that map $e_n \to x_n \to Tx_n \to e_n$. That is, $STR : \ell^p(\mathbb{N}) \to \ell^{p'}(\mathbb{N})$ maps $e_n$ to $e_n$, for all $n$. But the map $\ell^p(\mathbb{N}) \ni e_n \longmapsto e_n \in \ell^{p'}(\mathbb{N})$ is not bounded, as there are elements in $\ell^{p'}(\mathbb{N})$ that are not bounded in $\ell^p(\mathbb{N})$ (Exercise 2.8.21). This contradiction shows that $Tx_n \to 0$ and so $T$ is compact.

When $\mathcal{X} = c_0$, we have no reflexivity so the argument above does not apply. In the case where $p' > 1$ we can consider the adjoint, $T^* : \ell^{p'}(\mathbb{N}) \to \ell^1(\mathbb{N})$ and so the first part of the proof gives that $T^*$ is compact; and then $T$ is compact by Theorem 9.6.12. It remains to consider the case where $\mathcal{X} = c_0$ and $p' = 1$; but we proved such case in Proposition 9.6.11. ☐

The condition $p' < p$ in Theorem 9.8.16 cannot be relaxed, as for $p' \geq p$ the inclusion $\ell^p(\mathbb{N}) \to \ell^{p'}(\mathbb{N})$ is a non-compact norm-one operator. Here is another significant consequence of Pitt's Theorem:

***9.8.17. Corollary.***

*The Banach spaces $\ell^p(\mathbb{N})$, $1 \leq p \leq \infty$ are mutually non-isomorphic, and none of them is isomorphic to $c_0$. More generally, if $X_0$ is an infinite-dimensional closed subspace of any of $c_0$ or $\ell^p(\mathbb{N})$, $1 \leq p < \infty$, then it is not isomorphic as a Banach space to any subspace of any other of those spaces.*

*Proof.* We immediately discard $\ell^\infty(\mathbb{N})$ as it is non-separable while the rest are separable.

Because balls in an infinite-dimensional normed space are not relatively compact (Theorem 5.2.9) and bounded linear isomorphisms of closed subspaces are homeomorphisms by the Inverse Mapping Theorem (6.3.6), any such isomorphism would have to be non-compact. As we can always consider (by taking adjoints if necessary) the direction of the possible isomorphism to be $\ell^p(\mathbb{N}) \to \ell^{p'}(\mathbb{N})$ in which $p' < p$, we can apply Pitt's Theorem to conclude. □

The part about subspaces in Corollary 9.8.17 does not apply to $\ell^\infty(\mathbb{N})$, as it has $c_0$ as a closed subspace.

**9.8.18. Corollary.**

*Let $p \in (1, \infty)$. Then $\mathcal{K}(\ell^p(\mathbb{N}))$ is the unique proper (closed, two-sided) ideal of $\mathcal{B}(\ell^p(\mathbb{N}))$.*

*Proof.* Let $\mathcal{J} \subset \mathcal{B}(\ell^p(\mathbb{N}))$ be an ideal and let $T \in \mathcal{J} \setminus \mathcal{K}(\ell^p(\mathbb{N}))$. The middle paragraph of the proof Theorem 9.8.16 shows that if we take $p' = p$ there exist closed subspaces $\mathcal{X}_0, \mathcal{X}_1 \subset \ell^p(\mathbb{N})$ and bounded invertible maps

$$\ell^p(\mathbb{N}) \xrightarrow{R} \mathcal{X}_0 \xrightarrow{T} \mathcal{X}_1 \xrightarrow{S} \ell^p(\mathbb{N})$$

with $STR = I_{\ell^p(\mathbb{N})}$. As $\mathcal{X}_1$ is complemented, we can extend $S$ to a $\tilde{S} \in \mathcal{B}(\ell^p(\mathbb{N}))$ by making $\tilde{S} = SP$, where $P$ is a bounded projection onto $\mathcal{X}_1$. So now $R, \tilde{S} \in \mathcal{B}(\ell^p(\mathbb{N}))$ and $I_{\ell^p(\mathbb{N})} = \tilde{S}TR \in \mathcal{J}$; hence $\mathcal{J} = \mathcal{B}(\ell^p(\mathbb{N}))$. It follows that if $\mathcal{J}$ is a proper ideal, then $\mathcal{J} \subset \mathcal{K}(\ell^p(\mathbb{N}))$. Now, using Proposition 9.8.5 and Exercise 9.6.7,

$$\mathcal{K}(\ell^p(\mathbb{N})) = \overline{\mathcal{F}(\ell^p(\mathbb{N}))} \subset \mathcal{J} \subset \mathcal{K}(\ell^p(\mathbb{N})).$$

Hence $\mathcal{J} = \mathcal{K}(\ell^p(\mathbb{N}))$. □

**9.8.19. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. An operator $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ is* ***strictly singular*** *if for every closed subspace $\mathcal{X}_0 \subset \mathcal{X}$ with $\dim \mathcal{X}_0 = \infty$, $T|_{\mathcal{X}_0} : \mathcal{X}_0 \to \operatorname{ran} T$ is not bounded below. We will denote*

$$\mathcal{SS}(\mathcal{X}, \mathcal{Y}) = \{T \in \mathcal{B}(\mathcal{X}, \mathcal{Y}) : \ T \ \text{ is strictly singular}\}.$$

*When $\mathcal{Y} = \mathcal{X}$, we write $\mathcal{SS}(\mathcal{X})$.*

A trivial example of a strictly singular operator is a bounded linear functional, which can only be injective on a specific one-dimensional subspace. More generally, any compact operator on a Banach space is strictly singular (Exercise 9.6.9). We always have that $\mathcal{SS}(\mathcal{X})$ is a norm-closed double-sided ideal of $\mathcal{B}(\mathcal{X})$ and $\mathcal{K}(\mathcal{X}) \subset \mathcal{SS}(\mathcal{X})$ (Exercise 9.8.15).

***9.8.20. Corollary.***

*If $1 \leq p < p'$, then $\mathcal{B}(\ell^p(\mathbb{N}), \ell^{p'}(\mathbb{N})) = \mathcal{SS}(\ell^p(\mathbb{N}), \ell^{p'}(\mathbb{N}))$.*

*Proof.* Let $T \in \mathcal{B}(\ell^p(\mathbb{N}), \ell^{p'}(\mathbb{N}))$. By Theorem 9.8.16, $T$ is compact. Then $T$ is strictly singular by Exercise 9.6.9. $\square$

**9.8.3. Complemented Subspaces of $c_0$ and $\ell^p(\mathbb{N})$.** Next we delve a bit into complemented subspaces. Most results below are due to Pełczyński [Pel60]. The first one shows how strong Proposition 9.8.15 is. It is important to emphasize that when we say, in this context, that two subspaces are isomorphic, or that one "embeds" into the other, we require our isomorphisms to be bicontinuous (though not necessarily isometric).

***9.8.21. Proposition.***

*Let $\mathcal{X}$ be $\ell^p(\mathbb{N})$, $1 \leq p < \infty$, or $c_0$. If $\mathcal{X}_0 \subset \mathcal{X}$ is an infinite-dimensional closed subspace, then there exists a closed complemented (in $\mathcal{X}$) subspace $\mathcal{X}_1 \subset \mathcal{X}_0$ with $\mathcal{X}_1 \simeq \mathcal{X}$.*

*In particular, no infinite-dimensional reflexive space embeds in $c_0$ or $\ell^1(\mathbb{N})$.*

*Proof.* For each $n$ we can choose $x_n \in \mathcal{X}_0$ with $\|x_n\| = 1$ and $e_k^*(x_n) = 0$ for $k = 1, \ldots, n$. Indeed, the projection $P_n(\sum_{k=1}^\infty b_k e_k) = \sum_{k=1}^n b_k e_k$ cannot be injective on $\mathcal{X}_0$ since it has finite rank while $\dim \mathcal{X}_0 = \infty$. Then Proposition 9.8.15 applies to give us a subsequence $\{x_{n_k}\}$ that is a basic sequence, equivalent to the canonical basis of $\mathcal{X}$, and such that $\mathcal{X}_1 = \overline{\operatorname{span}}^{\|\cdot\|} \{x_{n_k}\}$ is complemented in $\mathcal{X}$. The equivalence with the canonical basis gives us $\mathcal{X}_1 \simeq \mathcal{X}$ via Proposition 9.8.9.

For the last assertion, we know that reflexivity passes to closed subspaces by Proposition 7.2.16. So if $\mathcal{Y}$ embeds in $c_0$ or $\ell^1(\mathbb{N})$, by the above they would have a reflexive subspace isomorphic to $c_0$ or $\ell^1(\mathbb{N})$, which are not reflexive. $\square$

It is trivial to "put" $\ell^1(\mathbb{N})$ into $c_0$, since every $x \in \ell^1(\mathbb{N})$ satisfies $\lim_n x(n) = 0$. But the identity map $\mathrm{id} : \ell^1(\mathbb{N}) \to c_0$, while bounded, is not bicontinuous. Indeed, it is easy to find, given $k > 0$, an element $x \in \ell^1(\mathbb{N})$ such that $\|x\|_\infty = 1$ and $\|x\|_1 > k$. What Proposition 9.8.21 says in that case is that there is no way to map $\ell^1(\mathbb{N})$ into $c_0$ in a bicontinuous way.

The following is a particularly remarkable result by Pełczyński, which opens the door for many others.

**9.8.22. Proposition. (Pełczyński Decomposition)**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces such that there exist linear injections $i : \mathcal{X} \to \mathcal{Y}$ and $j : \mathcal{Y} \to \mathcal{X}$ with $i(\mathcal{X})$ and $j(\mathcal{X})$ complementable in $\mathcal{Y}$ and $\mathcal{X}$ respectively. Assume moreover that either*

*(i) $\mathcal{X} \simeq \mathcal{X} \oplus \mathcal{X}$, $\mathcal{Y} \simeq \mathcal{Y} \oplus \mathcal{Y}$, or*

*(ii) $\mathcal{X} \simeq \bigoplus_{\mathbb{N}} \mathcal{X}$, where the direct sum is $c_0$, or $\ell^p$.*

*Then $\mathcal{X} \simeq \mathcal{Y}$.*

*Proof.* Note that $i(\mathcal{X})$ complementable implies that it is closed, so $i(\mathcal{X}) \simeq \mathcal{X}$.
(i) If $\mathcal{Z}$ is a complement of $i(\mathcal{X})$ in $\mathcal{Y}$, we have $\mathcal{Y} \simeq \mathcal{X} \oplus \mathcal{Z}$ by Proposition 5.3.5, and then

$$\mathcal{Y} \simeq \mathcal{X} \oplus \mathcal{Z} \simeq (\mathcal{X} \oplus \mathcal{X}) \oplus \mathcal{Z} \simeq \mathcal{X} \oplus (\mathcal{X} \oplus \mathcal{Z}) \simeq \mathcal{X} \oplus \mathcal{Y}.$$

Similarly,

$$\mathcal{X} \simeq \mathcal{Y} \oplus \mathcal{W} \simeq \mathcal{Y} \oplus \mathcal{Y} \oplus \mathcal{W} \simeq \mathcal{Y} \oplus \mathcal{X}.$$

As $\mathcal{X} \oplus \mathcal{Y} \simeq \mathcal{Y} \oplus \mathcal{X}$, we have obtained that $\mathcal{X} \simeq \mathcal{Y}$.
(ii) We still have $\mathcal{X} \simeq \mathcal{X} \oplus \mathcal{X}$ (Exercise 5.3.7). Also,

$$\mathcal{Y} \simeq \mathcal{X} \oplus \mathcal{Z} \simeq \mathcal{X} \oplus \mathcal{X} \oplus \mathcal{Z} \simeq \mathcal{X} \oplus \mathcal{Y}.$$

Then

$$\begin{aligned}\mathcal{X} &\simeq \bigoplus_n \mathcal{X} \simeq \Big(\bigoplus_n \mathcal{Y}\Big) \oplus \Big(\bigoplus_n \mathcal{W}\Big) \simeq \mathcal{Y} \oplus \Big(\bigoplus_n \mathcal{Y}\Big) \oplus \Big(\bigoplus_n \mathcal{W}\Big) \\ &\simeq \mathcal{Y} \oplus \Big(\bigoplus_n \mathcal{Y} \oplus \mathcal{W}\Big) \simeq \mathcal{Y} \oplus \bigoplus_n \mathcal{X} \simeq \mathcal{Y} \oplus \mathcal{X} \simeq \mathcal{Y}.\end{aligned}$$ □

Proposition 9.8.22 seems contrived, but we will immediately find it "in the wild" in the following amazing improvement over Proposition 9.8.21.

**9.8.23. Theorem. (Pełczyński)**

*Let $\mathcal{X}$ be $c_0$ or $\ell^p(\mathbb{N})$, $1 \le p < \infty$. If $\mathcal{X}_0 \subset \mathcal{X}$ is a complemented subspace with $\dim \mathcal{X} = \infty$, then $\mathcal{X}_0 \simeq \mathcal{X}$.*

*Proof.* By Proposition 9.8.21 there exists $\mathcal{X}_1 \subset \mathcal{X}_0$, complemented in $\mathcal{X}$ and isomorphic to $\mathcal{X}$. Because $\mathcal{X}_0$ is complemented in $\mathcal{X}$, we get that $\mathcal{X}_1$ is complemented in $\mathcal{X}_0$ (namely, if $\mathcal{X}_1 = P_1\mathcal{X}$, then $P_1\mathcal{X}_0 = \mathcal{X}_1$). So $\mathcal{X} \simeq \mathcal{X}_1 \subset \mathcal{X}_0$, and of course $\mathcal{X}_0 \subset \mathcal{X}$, complementable in both cases. Since $\ell^p(\mathbb{N}) \simeq \ell^p(\mathbb{N}) \oplus \ell^p(\mathbb{N})$ and $c_0 \simeq c_0 \oplus c_0$ (Exercise 9.8.10), then Proposition 9.8.22 guarantees that $\mathcal{X}_0 \simeq \mathcal{X}$. □

So in a sense the complementable subspaces of the $\ell^p(\mathbb{N})$ and $c_0$ spaces are really boring—an amazing discovery from Pełczyński—or we could say that these spaces are profoundly regular in the sense of the structure of their complementable subspaces. But we should not forget that this is about *complementable* subspaces. It turns out that every $\ell^p(\mathbb{N})$, $1 \leq p < \infty$ has a subspace without a Schauder basis, which of course is not complementable; this was proven by Davie [Dav73] and Szankowski [Sza78], following on Enflo's footsteps [Enf73].

More humbly, with the information at hand we can (easily!) prove the following.

**9.8.24. Corollary.**

*The Banach space $\ell^1(\mathbb{N})$ has a non-complemented closed subspace.*

*Proof.* Let $\mathcal{X}$ be $c_0$ (we can also take $\mathcal{X} = \ell^p(\mathbb{N})$ for $1 < p < \infty$). By Theorem 6.1.10 there exists a bounded linear surjection $Q : \ell^1(\mathbb{N}) \to \mathcal{X}$. Let $\mathcal{K} = \ker Q$, a closed subspace of $\ell^1(\mathbb{N})$. If $\mathcal{K}$ were complemented then $\ell^1(\mathbb{N}) = \mathcal{K} \oplus \mathcal{Z}$ for some closed subspace $\mathcal{Z}$, and we would have $\mathcal{X} \simeq \ell^1(\mathbb{N})/\mathcal{K} \simeq \mathcal{Z}$. By Theorem 9.8.23,

$$\ell^1(\mathbb{N}) \simeq Z \simeq \mathcal{X},$$

contradicting Corollary 9.8.17. So $\mathcal{K}$ is not complemented. $\square$

**9.8.4. Unconditional Convergence.** Traditionally,

$$\text{a series } \sum_{n=1}^{\infty} x_n \text{ converges if } \lim_{m\to\infty} \sum_{n=1}^{m} x_n \text{ exists.}$$

Meanwhile, recall that in a Banach space (duplicating a bit what we did in Definition 4.3.10)

**9.8.25. Definition.**

*a series $\sum_{j\in J} x_j$* ***converges unconditionally*** *to $x$ if for any $\varepsilon > 0$ there exists $F_0 \subset J$, finite, such that for all $F \supset F_0$, finite,*

$$\Big\| x - \sum_{j\in F} x_j \Big\| < \varepsilon. \tag{9.42}$$

The set $J$ is not required to be countable, nor to have any kind of order. If $\sum_k x_k$ and $\sum_k y_k$ converge unconditionally, then so does $\sum_k (a x_k + y_k)$ for $a \in \mathbb{C}$ (Exercise 9.8.17). The definition of unconditional convergence naturally implies the equivalent notion of "tails are small" that one expects from a convergent series. Concretely,

**9.8.26. Lemma.**

*Let $\mathcal{X}$ be a Banach space, $J$ a set, and $\{x_j\}_{j\in J} \subset \mathcal{X}$. Then the series $\sum_j x_j$ converges unconditionally if and only if for every $\varepsilon > 0$ there exists a finite set $F_0 \subset J$ such that for every finite $F_1 \subset J \setminus F$,*

$$\Big\| \sum_{j\in F_1} x_j \Big\| < \varepsilon.$$

*Proof.* If the series converges unconditionally to $x$, given $\varepsilon > 0$ we choose $F_\varepsilon$ such that (9.42) holds with $\varepsilon/2$. Then, for any $F_1 \subset J \setminus F$, the set $F_\varepsilon \cup F_1$ contains $F_\varepsilon$ and so

$$\frac{\varepsilon}{2} > \Big\| x - \sum_{j\in F_\varepsilon \cup F_1} x_j \Big\| \geq \Big\| \sum_{j\in F_1} x_j \Big\| - \Big\| x - \sum_{j\in F_\varepsilon} x_j \Big\| \geq \Big\| \sum_{j\in F_1} x_j \Big\| - \frac{\varepsilon}{2},$$

giving us that

$$\Big\| \sum_{j\in F_1} x_j \Big\| < \varepsilon.$$

Conversely, suppose that the tails are arbitrarily small. First we need to construct $x$. Given $\varepsilon > 0$, by hypothesis we have $F_\varepsilon$ such that the tails are small on its complement, up to $\varepsilon/2$. We want to say that the net $\{x_F\}$ indexed by the finite subsets $F \subset J$, and ordered by inclusion of said finite sets, is convergent. For any $F_1, F_2$ such that $F_1 \supset F_\varepsilon$ and $F_2 \supset F_\varepsilon$, we have

$$\begin{aligned}\Big\| \sum_{j\in F_1} x_j - \sum_{j\in F_2} x_j \Big\| &= \Big\| \sum_{j\in F_1\setminus F_\varepsilon} x_j - \sum_{j\in F_2\setminus F_\varepsilon} x_j \Big\| \\ &\leq \Big\| \sum_{j\in F_1\setminus F_\varepsilon} x_j \Big\| + \Big\| \sum_{j\in F_2\setminus F_\varepsilon} x_j \Big\| < \varepsilon.\end{aligned}$$

So the partial sums are Cauchy and thus a limit $x$ exists since $\mathcal{X}$ is Banach. $\square$

In the familiar case of the real Banach space $\mathbb{R}$, unconditional convergence is not the usual notion of convergence used in real analysis. Since we will be considering mostly countable series, let us recap. The naive definition of convergence of series is $\lim_{n\to\infty} \sum_{j=1}^{n} x_j$, and this is the usual meaning when unqualified. The problem with this definition, that the reader should know from real analysis classes, is conditional convergence; and that is the kind of pathology one tries to avoid with unconditional convergence. The following result should unravel unconditional convergence a bit (or, maybe, create more wonder):

**9.8.27. Proposition.**

*Let $\mathcal{X}$ be a Banach space, and $\{x_n\} \subset \mathcal{X}$ a sequence. Then the following statements are equivalent:*

*(i)* $\displaystyle\sum_{n=1}^{\infty} x_n$ *converges unconditionally;*

*(ii) for any* $b \in \ell^\infty(\mathbb{N})$*, the series* $\displaystyle\sum_{k=1}^{\infty} b_k x_k$ *converges;*

*(iii) for any* $b \in \{1,-1\}^{\mathbb{N}}$*, the series* $\displaystyle\sum_{k=1}^{\infty} b_k x_k$ *converges;*

*(iv) there exists* $L \subset \mathbb{C}$*, bounded, with* $|L| \geq 2$*, such that the series* $\displaystyle\sum_{k=1}^{\infty} b_k x_k$ *converges for any* $b \in L^{\mathbb{N}}$*;*

*(v)* $\displaystyle\sum_{k=1}^{\infty} x_{n_k}$ *converges for every increasing sequence* $\{n_k\}_k \subset \mathbb{N}$*;*

*(vi)* $\displaystyle\sum_{n=1}^{\infty} x_{\sigma(n)}$ *converges for every permutation* $\sigma$ *of* $\mathbb{N}$*.*

*Proof.* (i) $\Longrightarrow$ (ii) It is enough to prove the case where $b_k \in \mathbb{R}$ for all $k$, since

$$\sum_{k=1}^{\infty} b_k x_k = \sum_{k=1}^{\infty} (\operatorname{Re} b_k)\, x_k + i \sum_{k=1}^{\infty} (\operatorname{Im} b_k)\, x_k$$

and the left-hand-side will converge if both series in the right-hand-side converge. So we assume that $b_k \in \mathbb{R}$ for all $k$. Since multiplying everything by a fixed nonzero scalar will not change convergence, we may also assume that $|b_k| \leq 1$ for all $k$.

The key fact to prove the real case is this equality: given $x_1, \ldots, x_n \in \mathcal{X}$,

$$\begin{aligned}
\max&\left\{ \left\| \sum_{k=1}^{n} b_k x_k \right\| : \; b_k \in [-1,1] \right\} \\
&= \max\left\{ \sum_{k=1}^{n} |\operatorname{Re} g(x_k)| : \; g \in \mathcal{X}^*, \; \|g\| = 1 \right\} \qquad (9.43)\\
&= \max\left\{ \left\| \sum_{k=1}^{n} b_k x_k \right\| : \; b_k = \pm 1 \right\}
\end{aligned}$$

(one would initially write sup instead of max, but all three sets are continuous images of compact domains). To prove (9.43), let $g \in \mathcal{X}^*$ with $\|g\| = 1$ and $g\left(\sum_{k=1}^{n} b_k x_k\right) = \left\|\sum_{k=1}^{n} b_k x_k\right\|$ (cfr. Corollary 5.7.7). Then

$$\left\|\sum_{k=1}^{n} b_k x_k\right\| = g\left(\sum_{k=1}^{n} b_k x_k\right) = \operatorname{Re} g\left(\sum_{k=1}^{n} b_k x_k\right) = \sum_{k=1}^{n} b_k \operatorname{Re} g(x_k)$$

$$\leq \sum_{k=1}^{n} |b_k|\, |\operatorname{Re} g(x_k)| \leq \sum_{k=1}^{n} |\operatorname{Re} g(x_k)|.$$

Conversely, given $g \in \mathcal{X}^*$ with $\|g\| = 1$ choose $b_k \in \{1, -1\}$ with $b_k \operatorname{Re} g(x_k) = |\operatorname{Re} g(x_k)|$. Then

$$\sum_{k=1}^{n} |\operatorname{Re} g(x_k)| = \sum_{k=1}^{n} b_k \operatorname{Re} g(x_k) = \operatorname{Re} g\left(\sum_{k=1}^{n} b_k x_k\right) \leq \left\|\sum_{k=1}^{n} b_k x_k\right\|,$$

completing the proof of the first line of (9.43). The equality with the last line follows easily from the fact that in the proof of the first line we chose all $b_k$ with $b_k = \pm 1$. We note also that Lemma 9.8.26 implies that $\lim_n \|x_n\| = 0$. Now, given $\varepsilon > 0$, by the unconditional convergence there exists $F_0$, finite, such that $\left\|\sum_{k \in F_1} x_k\right\| < \varepsilon$ for all finite $F_1 \subset \mathbb{N} \setminus F_0$. Given coefficients $a_k \in \{1, -1\}$, if we put $F_1^+ = \{k : \ a_k = 1\}$, $F_1^- = \{k : \ a_k = -1\}$, then

$$\left\|\sum_{k \in F_1} a_k x_k\right\| = \left\|\sum_{k \in F_1^+} x_k - \sum_{k \in F_1^-} x_k\right\| \leq \left\|\sum_{k \in F_1^+} x_k\right\| + \left\|\sum_{k \in F_1^-} x_k\right\| < 2\varepsilon.$$

By (9.43), if $b_k \in [-1, 1]$,

$$\left\|\sum_{k \in F_1} b_k x_k\right\| \leq \max\left\{\left\|\sum_{k \in F_1} a_k x_k\right\| : \ a_k = \pm 1\right\} < 2\varepsilon.$$

So $\sum_{k=1}^{\infty} b_k x_k$ converges.

(ii) $\iff$ (iii) this follows from (9.43) and the fact that coefficients can be split in their real and imaginary parts.

(iii) $\implies$ (iv) Take $L = \{1, -1\}$.

(iv) $\implies$ (v) Let $r_1, r_2 \in L$, with $r_1 \neq r_2$ . Let $G = \{n_k\}$, and $H = \mathbb{N} \setminus G$. By hypothesis the series

$$\sum_k r_1 x_k, \qquad \text{and} \qquad \sum_k b_k x_k$$

converge, where
$$b_k = \begin{cases} r_1, & k \in H \\ r_2, & k \in G \end{cases}$$
Then the difference converges, and this is
$$\sum_k r_1 x_k - \sum_k b_k x_k = \sum_{k\in G}(r_1 - r_2)x_k = (r_1 - r_2)\sum_k x_{n_k}.$$
Hence $\sum_k x_{n_k}$ converges.

(v) $\Longrightarrow$ (i) Suppose that the series does not converge unconditionally. Then, using Lemma 9.8.26 there exists $\varepsilon > 0$ and finite sets $\{F_k\}$ in $\mathbb{N}$ such that $\max F_k < \min F_{k+1}$ for all $k$, and with
$$\left\| \sum_{n\in F_k} x_k \right\| \geq \varepsilon.$$
If we take $\{n_k\}$ to be an increasing enumeration of $\bigcup_k F_k$, then for any $m \in \mathbb{N}$ we can choose $F_k$ with $\min F_k > m$. This gives
$$\left\| \sum_{n\in F_k\cup F_{k+1}} x_n - \sum_{n\in F_k} x_n \right\| = \left\| \sum_{n\in F_{k+1}} x_n \right\| \geq \varepsilon$$
and the partial sums of the series $\sum_k x_{n_k}$ are not Cauchy, a contradiction.

(i) $\Longrightarrow$ (vi) Fix a permutation $\sigma$. Fix $\varepsilon > 0$ and let $F_0 \subset \mathbb{N}$, finite, as in the unconditional convergence hypothesis. Write $x$ for the limit of the series. Let $m = \max \sigma^{-1}(F_0)$. Then, for any $m' \geq m$, $\sigma^{-1}(F_0) \subset \{1, \ldots, m'\}$, so $F_0 \subset \{\sigma(1), \ldots, \sigma(m')\}$. Thus for any $m' > m$ we have
$$\left\| x - \sum_{n=1}^{m'} x_{\sigma(n)} \right\| < \varepsilon,$$
showing that $\sum_{n=1}^{\infty} x_{\sigma(n)}$ converges.

(vi) $\Longrightarrow$ (i) Suppose that the series does not converge unconditionally. Then, using Lemma 9.8.26 there exists $\varepsilon > 0$ and finite sets $\{F_k\}_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that $\max F_k < \min F_{k+1}$ for all $k$, and with
$$\left\| \sum_{n\in F_k} x_k \right\| \geq \varepsilon, \qquad\qquad k \in \mathbb{N}.$$
If we put $E_0 = \{0\}$ and $E_n = \{\max F_{n-1} + 1, \ldots, \max F_n\} \setminus F_n$, then
$$E_1, F_1, E_2, F_2, \ldots$$
covers all of $\mathbb{N}$. By taking $\sigma$ to be the permutation associated with this order, we have that $\sum_k x_{\sigma(k)}$ is not Cauchy, since for arbitrarily big $k$ we will find chunks with norm greater than $\varepsilon$. $\square$

When the Banach space $\mathcal{X}$ is finite-dimensional, the conditions in Proposition 9.8.27 are also equivalent to **absolute convergence**, i.e. the case where $\sum_j \|x_j\| < \infty$ (Exercises 9.8.18 and 9.8.19). On an infinite-dimensional Banach space, though, absolute convergence is not equivalent to unconditional convergence, but actually stronger. And unconditional convergence is enough to guarantee existence of the series in a useful way. For an obvious example, consider the element $\xi = \sum_n \frac{1}{n}\,\xi_n$ in a separable Hilbert space with orthonormal basis $\{\xi_n\}$; then the series converges unconditionally but not absolutely. The exact same example gives an unconditionally but not absolutely convergent series in $\ell^\infty(\mathbb{N})$. It was shown in [DR50a] that if in a Banach space $\mathcal{X}$ every unconditionally convergent series converges absolutely, the $\mathcal{X}$ is finite-dimensional.

***9.8.28. Corollary.***

*Let $\mathcal{X}$ be a Banach space and $\{x_n\} \subset \mathcal{X}$ a sequence such that $\sum_n x_n$ converges unconditionally. Then $\sum_n b_n x_n$ converges uniformly for $b$ on any bounded subset of $\ell^\infty(\mathbb{N})$. Concretely, given $\varepsilon > 0$ there exists $F_0 \subset \mathbb{N}$, finite, such that for any $b \in \ell^\infty(\mathbb{N})$ and any $F_1 \subset \mathbb{N} \setminus F_0$ we have*

$$\left\| \sum_{n\in F_1} b_n x_n \right\| \leq \varepsilon \, \|b\|_\infty .$$

*Proof.* Suppose first that $\|b\|_\infty = 1$. Fix $\varepsilon > 0$. By the unconditional convergence there exists $F_0 \subset \mathbb{N}$, finite, such that for any $F_1 \subset \mathbb{N} \setminus F_0$ finite we have

$$\left\| \sum_{n\in F_1} x_n \right\| < \frac{\varepsilon}{4}.$$

If $a \in \{1,-1\}^{\mathbb{N}}$ we put $F_1^+ = \{n \in F_1 : \ a_n = 1\}$, and $F_1^- = \{n \in F_1 : \ a_n = -1\}$, and then

$$\left\| \sum_{n\in F_1} a_n x_n \right\| = \left\| \sum_{n\in F_1^+} x_n - \sum_{n\in F_1^-} x_n \right\| \leq \left\| \sum_{n\in F_1^+} x_n \right\| + \left\| \sum_{n\in F_1^-} x_n \right\| < \frac{2\varepsilon}{4} = \frac{\varepsilon}{2}.$$

Using (9.43),

$$\begin{aligned}\left\| \sum_{n\in F_1} b_n x_n \right\| &\leq \left\| \sum_{n\in F_1} (\operatorname{Re} b_n) x_n \right\| + \left\| \sum_{n\in F_1} (\operatorname{Im} b_n) x_n \right\| \\ &\leq 2 \max \left\{ \left\| \sum_{n\in F_1} a_n x_n \right\| : \ a \in \{1,-1\}^{\mathbb{N}} \right\} < \frac{2\varepsilon}{2} = \varepsilon.\end{aligned}$$

For arbitrary $b$, we simply scale with $\|b\|_\infty$. $\square$

**9.8.5. Bounded Operators on $c_0$.** The next result will be crucial in proving several results concerning operators on $c_0$.

***9.8.29. Proposition.***

*Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(c_0, \mathcal{X})$. Denote as usual the canonical basis by $\{e_k\}$. Then the following statements are equivalent:*

*(i) $T$ is compact;*

*(ii) $\sum_{k=1}^{\infty} Te_k$ converges unconditionally.*

*Proof.* (i) $\implies$ (ii) The double adjoint of $T$ is $T^{**} : \ell^\infty(\mathbb{N}) \to \mathcal{X}^{**}$. By Proposition 9.6.5, $T^{**}$ maps weak*-convergent sequences to norm-convergent sequences. By Corollary 7.2.20 the closed unit ball of $\ell^\infty(\mathbb{N})$ (as the dual of $\ell^1(\mathbb{N})$) is metrizable in the weak*-topology. So, on the closed unit ball $\overline{B_1^{\ell^\infty}(0)}$, $T^{**}$ is weak*-norm continuous, since the metrizability allows us to use sequences instead of nets. For any permutation $\sigma$ of $\mathbb{N}$, the series $\sum_{k=1}^{\infty} e_{\sigma(k)}$ converges weak* to 1 in $\overline{B_1^{\ell^\infty}(0)}$. Thus $\sum_{k=1}^{\infty} T^{**}e_{\sigma(k)}$ converges in norm in $\mathcal{X}^{**}$. As $T^{**}e_k = Te_k$ for all $k$ and the norm topology on $\mathcal{X}^{**}$ agrees with the norm topology on $\mathcal{X}$, we get that $\sum_{k=1}^{\infty} Te_{\sigma(k)}$ converges in $\mathcal{X}$. This works for any permutation $\sigma$, so by Proposition 9.8.27 the series $\sum_{k=1}^{\infty} Te_k$ converges unconditionally.

(ii) $\implies$ (i) Assume without loss of generality that $\|T\| = 1$. Denote by $P_n : c_0 \to c_0$ the projection onto $\operatorname{span}\{e_1 \dots, e_n\}$. Let $x \in B_1^{c_0}(0)$ and fix $\varepsilon > 0$. We have $x = \sum_k b_k e_k$, with $\|b\|_\infty = \|x\|_\infty$. By the unconditional convergence hypothesis and Corollary 9.8.28, there exists $F_0 \subset \mathbb{N}$, finite, such that

$$\left\| \sum_{k \in F_1} b_k Te_k \right\| < \varepsilon \|b\|_\infty$$

for all finite subsets $F_1 \subset \mathbb{N} \setminus F_0$. Let $n_0 = \max F_0$. Then

$$\|(T - TP_{n_0})x\| = \left\| \sum_{k > n_0} b_k Te_k \right\| = \lim_{F \in n_0 + \mathbb{N}} \left\| \sum_{k \in F} b_k Te_k \right\| \le \varepsilon \|b\|_\infty = \varepsilon \|x\|_\infty.$$

So $T = \lim_n TP_n \in \mathcal{K}(c_0, \mathcal{X})$ by Proposition 9.6.2. $\square$

We now get some consequences out of Proposition 9.8.29. First one is

***9.8.30. Proposition.***

*Let $\mathcal{X}$ be a Banach space. Then $\mathcal{SS}(c_0, \mathcal{X}) = \mathcal{K}(c_0, \mathcal{X})$.*

*Proof.* Recall that $\mathcal{K}(\mathcal{X}) \subset \mathcal{SS}(\mathcal{X})$. Suppose that $T$ is not compact. By Proposition 9.8.29, the series $\sum_k Te_k$ does not converge unconditionally. Then by Lemma 9.8.26 there exist $\varepsilon > 0$ and disjoint finite subsets $F_n \subset \mathbb{N}$, $n \in \mathbb{N}$, with $\left\| \sum_{k \in F_n} Te_k \right\| \geq \varepsilon$. Given $\varphi \in c_0^* = \ell^1(\mathbb{N})$, say $\varphi = y$,

$$\varphi\Big( \sum_{k \in F_n} e_k \Big) = \sum_{k \in F_n} y_k \xrightarrow[n \to \infty]{} 0$$

since for $n$ big enough the sets $F_n$ will only contain big enough integers. As $\varphi \circ T \in \mathcal{X}^*$, this implies that $x_n \xrightarrow{\text{weak}} 0$, where

$$x_n = \sum_{k \in F_n} Te_k.$$

Then Theorem 9.8.12 shows that there exists a subsequence $X = \{x_{n_j}\}$ which is a basic sequence; we still denote this subsequence as $\{x_n\}$ to simplify notation. Given $c \in c_{00}$,

$$\left\| \sum_{n=1}^{\infty} c_n x_n \right\| = \left\| T\Big( \sum_{n=1}^{\infty} c_n \sum_{k \in F_n} e_k \Big) \right\| \leq \|T\| \, \|c\|_\infty \tag{9.44}$$

as the $F_n$ are disjoint. We also have, if $\{Q_n\}$ denote the basis projections of $\{x_n\}$,

$$\begin{aligned} |c_m| = \|x_m\|^{-1} \|c_m x_m\| &= \|x_m\|^{-1} \left\| (Q_m - Q_{m-1}) \Big( \sum_{n=1}^{\infty} c_n x_n \Big) \right\| \\ &\leq \varepsilon^{-1} \, 2b_X \left\| \sum_{n=1}^{\infty} c_n x_n \right\|. \end{aligned} \tag{9.45}$$

As this bound works for all $m$, it is a bound for $\|c\|_\infty$. Now (9.44) and (9.45) combined show that $\{x_n\}$ is equivalent to the canonical basis of $c_0$ and thus also equivalent to the block sequence $\big( \sum_{k \in F_n} e_k \big)_n$. By Proposition 9.8.9,

$$T|_{\overline{\text{span}}^{\|\cdot\|}\{x_n\}} : \overline{\text{span}}^{\|\cdot\|}\{x_n\} \to \overline{\text{span}}^{\|\cdot\|} \Big\{ \sum_{k \in F_n} e_k : \ n \Big\}$$

is invertible. This means that $T$ is bounded below on an infinite-dimensional subspace, and so not strictly singular. $\square$

And here is a second consequence of Proposition 9.8.29.

***9.8.31. Theorem. (Pełczyński)***

*Let $\mathcal{X}$ be a Banach space and $T \in \mathcal{B}(c_0, \mathcal{X})$ not compact. Then there exists a subspace $Z \subset c_0$ such that $Z \simeq c_0$ and $T|_Z$ is injective.*

*Proof.* We know from Proposition 9.8.29 that $\sum_{k=1}^{\infty} Te_k$ is not unconditionally convergent. By Lemma 9.8.26 there exists $\varepsilon > 0$ and disjoint finite sets $\{F_k\}_{k\in\mathbb{N}}$ such that

$$\Big\|\sum_{n\in F_k} Te_n\Big\| \geq \varepsilon. \tag{9.46}$$

For each $n \in \mathbb{N}$, let

$$x_n = \sum_{k\in F_n} e_k \in c_0, \qquad y_n = \sum_{k\in F_n} Te_k \in \mathcal{X}.$$

We have $x_n \xrightarrow{\text{weak}} 0$. As $T$ is bounded, $y_n = Tx_n \xrightarrow{\text{weak}} 0$. We have $\|y_n\| \geq \varepsilon$ for all $n$ by (9.46). Note that $\{x_n\}$ is a block basic sequence for the canonical basis $\{e_n\}$ in $c_0$, and so a basic sequence by Lemma 9.8.10. By the Bessaga-Pełczyński Selection Principle (Theorem 9.8.12), applied to the Schauder basis $\{x_n\}$ of $\overline{\operatorname{span}}^{\|\cdot\|}\{x_n\}$ and $\{y_n\}$, there exists a subsequence of $\{y_n\}$, that we will still denote by $\{y_n\}$, that is equivalent to a block basic sequence of $\{x_n\}$. So we have that, after possibly passing to subsequences and reframing the $x_n$ to be sums over bigger finite sets, $\{x_n\}$ and $\{y_n\}$ are equivalent basic sequences.

Next we claim that $\{y_n\}$ is equivalent to the canonical basis $\{e_n\}$ of $c_0$. For this, given $z \in c_{00}$ we have

$$\begin{aligned}\Big\|\sum_{n=1}^{\infty} z(n)\, y_n\Big\| &= \Big\|T\Big(\sum_{n=1}^{\infty} z(n)\, x_n\Big)\Big\| \leq \|T\| \,\Big\|\sum_{n=1}^{\infty}\sum_{k\in F_n} z(n)\, e_k\Big\| \\ &\leq \|T\| \max\{|z(n)| :\ n\} = \|T\|\,\|z\|_\infty.\end{aligned}$$

Conversely, since $Y = \{y_n\}$ is basic with $\|y_n\| \geq \varepsilon$ for all $n$, choosing $n_0$ such that $\|z\|_\infty = |z(n_0)|$, and writing $P_n$ for the basis projections for $\{y_n\}$,

$$\begin{aligned}\|z\|_\infty = |z(n_0)| &= \|y_{n_0}\|^{-1}\Big\|(P_{n_0} - P_{n_0-1})\Big(\sum_{n=1}^{\infty} z(n)\, y_n\Big)\Big\| \\ &\leq 2\varepsilon^{-1}\, \mathrm{b}_Y \,\Big\|\sum_{n=1}^{\infty} z(n)\, y_n\Big\|.\end{aligned}$$

Thus, by the density of $c_{00}$ in $c_0$, the map $S : c_0 \to \overline{\operatorname{span}}^{\|\cdot\|}\{y_n\}$ given by $Sz = \sum_{n=1}^{\infty} z(n)\, y_n$ is invertible and so $\{e_n\}$ and $\{y_n\}$ are equivalent. As $\{x_n\}$

is equivalent to $\{y_n\}$, it is also equivalent to $\{e_n\}$. So if we take $Z \subset c_0$ given by $Z = \overline{\operatorname{span}}^{\|\cdot\|}\{x_n\}$, then $Z \simeq c_0$ and $T|_Z$ is injective, since

$$\left\| T\left(\sum_{n=1}^{\infty} z(n)\, x_n\right)\right\| = \left\| \sum_{n=1}^{\infty} z(n)\, y_n \right\| \geq \frac{\varepsilon}{2b_Y}\, \|z\|_\infty. \qquad \square$$

### *9.8.32. Corollary.*

*Let $\mathcal{X}$ be a Banach space. The following statements are equivalent:*

*(i) $\mathcal{X}$ does not contain a copy of $c_0$;*

*(ii) if $\{x_n\} \subset \mathcal{X}$ and the set*

$$\mathcal{X}_1 = \Big\{ \sum_{k=1}^{n} \delta_k\, x_k : \ \delta_k = \pm 1, \ n \in \mathbb{N} \Big\}$$

*is bounded, then $\mathcal{X}_1$ is relatively compact;*

*(iii) if $\{x_n\} \subset \mathcal{X}$ and the set*

$$\mathcal{X}_1 = \Big\{ \sum_{k=1}^{n} \delta_k\, x_k : \ \delta_k = \pm 1, \ n \in \mathbb{N} \Big\}$$

*is bounded, then $\sum_{n=1}^{\infty} x_n$ is unconditionally convergent.*

*Proof.* (i) $\implies$ (iii) Let $T : c_{00} \to \mathcal{X}$ be the linear operator induced by $Te_k = x_k$. By the boundedness of $\mathcal{X}_1$ and (9.43), $T$ is bounded and so it extends to $T \in \mathcal{B}(c_0, \mathcal{X})$. If $\sum_{n=1}^{\infty} x_n$ is not unconditionally convergent, then $T$ is not compact by Proposition 9.8.29; by Theorem 9.8.31, $\mathcal{X}$ contains a copy of $c_0$.

(iii) $\implies$ (ii) We are assuming that $\mathcal{X}_1$ is bounded, and that $\displaystyle\sum_{n=1}^{\infty} x_n$ is unconditionally convergent. Consider the Cantor space $2^{\mathbb{N}}$, expressed as $\{1, -1\}^{\mathbb{N}}$, with the product topology and define $S : 2^{\mathbb{N}} \to \mathcal{X}$ by $Sz = \sum_{n=1}^{\infty} z(n)\, x_n$. These series all converge by Proposition 9.8.27, so $S$ is well-defined. Next we show that $S$ is continuous. Fix $\varepsilon > 0$ and $z \in 2^{\mathbb{N}}$. By the unconditional convergence there exists $F_0$ finite such that $\displaystyle\Big\| \sum_{k\in F} x_k \Big\| < \varepsilon$ for all $F \in \mathbb{N} \setminus F_0$. Then for any $w$ in any neighbourhood $N_F = \{v \in 2^{\mathbb{N}} : |v(n) - z(n)| < \varepsilon, \ n \in F\}$ (where $F \supset F_0$)

$$\begin{aligned} \|Sw - Sz\| &= \Big\| \sum_{n\in F} (w(n) - z(n))\, x_n \Big\| \\ &= \Big\| \sum_{\{n\in F:\ w(n)=1\}} x_n \Big\| + \Big\| \sum_{\{n\in F:\ w(n)=-1\}} x_n \Big\| \end{aligned}$$

$$+\left\|\sum_{\{n\in F:\ z(n)=1\}} x_n\right\| + \left\|\sum_{\{n\in F:\ z(n)=-1\}} x_n\right\|$$
$$< 4\varepsilon.$$

Hence $S2^{\mathbb{N}}$ is compact, being a continuous image of a compact set, and $\mathcal{X}_1 \subset S2^{\mathbb{N}}$ is relatively compact.

(ii) $\Longrightarrow$ (i) Suppose that $\mathcal{X}$ contains a copy of $c_0$. Then we have a bounded injection $T : c_0 \to \mathcal{X}$ with closed range (isomorphic to $c_0$). The set

$$\mathcal{X}_2 = \left\{\sum_{k=1}^{n} \delta_k\, e_k :\ \delta_k = \pm 1,\ n \in \mathbb{N}\right\} \subset c_0$$

is not relatively compact, since it contains for instance $\left\{\sum_{k=1}^{n} e_k\right\}_n$, where all the elements are at distance 1 from each other. Because $T$ is bounded and injective with closed range, it is invertible onto its range by the Inverse Mapping Theorem (Theorem 6.3.6) and hence a homeomorphism. Thus $\mathcal{X}_1 = T\mathcal{X}_2$ is not relatively compact, though it is bounded (since $\mathcal{X}_2 \subset B_1^{c_0}(0)$ and $T$ is bounded). $\square$

Here is another peculiarity of $c_0$. The following is an extension of Sobczyck's Theorem [Sob41]. The proof below appeared in [Vee71].

***9.8.33. Proposition.***

*Let $\mathcal{X}$ be a separable Banach space , $\mathcal{X}_0 \subset \mathcal{X}$ a closed subspace, and $T \in \mathcal{B}(\mathcal{X}_0, c_0)$. Then there exists $\tilde{T} \in \mathcal{B}(\mathcal{X}, c_0)$, with $\tilde{T}|_{\mathcal{X}_0} = T$ and $\|\tilde{T}\| \leq 2\|T\|$.*

*Proof.* The closed ball $\overline{B_{\|T\|}^{\mathcal{X}^*}(0)}$ is compact and metrizable in the weak*-topology by Corollary 7.2.20. Let $d$ be such a metric. For each $n \in \mathbb{N}$, let $\varphi_n \in \mathcal{X}^*$ be given by $\varphi_n(x) = e_n^*(Tx)$ on $\mathcal{X}_0$ and extended by Hahn–Banach (Corollary 5.7.6) to all of $\mathcal{X}$. Put $K = \overline{B_{\|T\|}^{\mathcal{X}^*}(0)} \cap (\mathcal{X}_0)^o$. We have $\|\varphi_n\| = \|e_n^* \circ T\| \leq \|T\|$. If $\varphi$ is weak*-accumulation point of $\varphi_n$, then $\|\varphi\| \leq \|T\|$ and $\varphi|_{\mathcal{X}_0} = 0$ (by way of the range of $T$ being in $c_0$), so $\varphi \in K$. This means that $d(\varphi_n, K) \to 0$. So there exists $\{\psi_n\} \subset K$ such that $d(\varphi_n, \psi_n) \to 0$. In particular $\varphi_n(x) - \psi_n(x) \to 0$ for all $x \in \mathcal{X}$. Let $\tilde{T} : \mathcal{X} \to c_0$ be given by $(\tilde{T}x)(n) = \varphi_n(x) - \psi_n(x)$. We have $\|\tilde{T}\| \leq 2\|T\|$. When $x \in \mathcal{X}_0$, as $\psi_n \in K$ we have $\psi_n(x) = 0$ for all $n$; hence $(\tilde{T}x)(n) = \varphi_n(x) = (Tx)(n)$ for all $n$ and so $\tilde{T}x = Tx$. $\square$

A consequence of Proposition 9.8.33 is that for $c_0$ the converse of Theorem 9.8.23 holds:

***9.8.34. Corollary.***

*Let $\mathcal{X}_0 \subset c_0$ be a closed subspace. If $\mathcal{X}_0$ is isomorphic to $c_0$, then $\mathcal{X}_0$ is complemented.*

*Proof.* Let $\psi : \mathcal{X}_0 \to c_0$ be an isomorphism. By Proposition 9.8.33 there exists $\tilde{\psi} \in \mathcal{B}(c_0)$ such that $\tilde{\psi}|_{\mathcal{X}_0} = \psi$. Let $P \in \mathcal{B}(c_0)$ be $P = \psi^{-1} \circ \tilde{\psi}$. Then $Px \in \mathcal{X}_0$ for all $x \in c_0$ and $P^2 = P$. $\square$

**9.8.6. Exercises.**

(9.8.1) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that $P_n$ is linear and bounded.

(9.8.2) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that $b_E < \infty$.

(9.8.3) In the proof of Proposition 9.8.2, show that $\|\cdot\|_b$ is a seminorm.

(9.8.4) Let $\mathcal{X}$ be a Banach space and $E = \{e_n\}$ a Schauder basis. Show that $E$ is also a Schauder basis for the Banach space $(\mathcal{X}, \|\cdot\|_b\|)$.

(9.8.5) Fix $t_1 = 0$, $t_2 = 1$, and $t_3 < \cdots < t_n \in (0,1)$. Let $D_n \subset C[0,1]$ be set of piecewise linear functions with nodes at each $t_j$. Show that $D_n$ is a subspace and that $\dim D_n = n$.

(9.8.6) Let $\mathcal{X}$ be a Banach space and $\{e_n\}$ a sequence of nonzero elements satisfying (9.39). Show that $\{e_n\}$ is linearly independent.

(9.8.7) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $\mathcal{X}_0 \subset \mathcal{X}$, $\mathcal{Y}_0 \subset \mathcal{Y}$ (not necessarily closed) subspaces, and $T : \mathcal{X}_0 \to \mathcal{Y}_0$ a bounded linear invertible operator. Show that there exists a unique extension $\tilde{T} : \overline{\mathcal{X}_0} \to \overline{\mathcal{Y}_0}$ that is still invertible. In other words, a bounded isomorphism of subspaces extends to an isomorphism of their closures.

(9.8.8) Show that any subsequence of a basic sequence is a basic sequence.

(9.8.9) In the proof of (i) $\implies$ (ii) in Proposition 9.8.9, show that $S$ is bijective.

(9.8.10) Show that

(i) $c_0 \simeq c_0 \oplus_\infty c_0$, where $\|(x,y)\|_\infty = \max\{\|x\|_\infty, \|y\|_\infty\}$;

(ii) $\ell^p(\mathbb{N}) \simeq \ell^p(\mathbb{N}) \oplus_p \ell^p(\mathbb{N})$, $1 \le p \le \infty$, where

$$\|(x,y)\|_p = (\|x\|_p^p + \|y\|_p^p)^{1/p}.$$

In both cases the isomorphism can be made isometric.

(9.8.11) Show that

(i) $c_0 \simeq \left( \bigoplus_{n\in\mathbb{N}} c_0 \right)_{c_0}$, where $\left\| \bigoplus_n x_n \right\|_\infty = \max\{\|x_n\|_\infty : n\}$;

(ii) $\ell^p(\mathbb{N}) \simeq \bigoplus_{n\in\mathbb{N}} \ell^p(\mathbb{N})$, $1 \leq p \leq \infty$, where

$$\Big\| \bigoplus_n x_n \Big\|_p = \Big( \sum_{n=1}^\infty \|x_n\|_p^p \Big)^{1/p}.$$

In both cases the isomorphism can be made isometric. *The only difficulty of this exercise compared to Exercise 9.8.10 is having to deal with sequences of sequences and not getting swamped by the notation. If needed, it might help to use $x(k)$ for the $k^{\text{th}}$ entry of $x$.*

(9.8.12) Corollary 9.8.13 is a result due to Mazur, but the proof offered is not the original one. Here we will outline the original argument.

(i) Given a normed space $\mathcal{X}$ and $V \subset \mathcal{X}$ a finite-dimensional subspace, and given $\varepsilon \in (0,1)$, show that there exists $x \in \mathcal{X}$ with $\|x\| = 1$ and such that

$$\|z\| \leq (1+\varepsilon)\|z + \lambda x\|, \qquad z \in V,\ \lambda \in \mathbb{C}. \tag{9.47}$$

*(Hint: find an $\varepsilon/2$-net for $\partial B_1^V(0)$, bounded linear functionals that are $1$ at each point of the net, and take $x$ in the intersection of the kernels)*

(ii) Given $\varepsilon > 0$, show that there exists $\delta > 0$ such that

$$\prod_{k=1}^\infty (1 + \delta\, 2^{-k}) \leq 1 + \varepsilon.$$

(iii) Use Proposition 9.8.6 to prove Corollary 9.8.13, by using (i) inductively.

(9.8.13) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, $\mathcal{X}_0 \subset \mathcal{X}$ a subspace with $\dim \mathcal{X}_0 = \infty$, and $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$. Show that there exists a normalized basic sequence $X = \{x_n\} \subset \mathcal{X}_0$ with $b_X < 2$ and such that $\|Tx_n\| < 2^{-n}$ for all $n$.

(9.8.14) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$. Show that $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ if and only if for every infinite-dimensional subspace $\mathcal{X}_0 \subset \mathcal{X}$ there exists an infinite-dimensional subspace $\mathcal{X}_1 \subset \mathcal{X}_0$ with $T|_{\mathcal{X}_1}$ compact.

(9.8.15) Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces. Show that

(i) $\mathcal{SS}(\mathcal{X}, \mathcal{Y})$ is a norm-closed subspace;

(ii) if $T \in \mathcal{SS}(\mathcal{X}, \mathcal{Y})$ and $S \in \mathcal{B}(\mathcal{Y}, \mathcal{Z})$, $R \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$, then $ST \in \mathcal{SS}(\mathcal{X}, \mathcal{Z})$ and $TR \in \mathcal{SS}(\mathcal{Z}, \mathcal{X})$;

(iii) $\mathcal{K}(\mathcal{X}, \mathcal{Y}) \subset \mathcal{SS}(\mathcal{X}, \mathcal{Y})$.

(9.8.16) Show that $\mathcal{B}(c_0)$ has a unique non-trivial (closed, double-sided) ideal.

(9.8.17) Show that if $\sum_k x_k$ and $\sum_k y_k$ converge unconditionally, then so does $\sum_k (ax_k + y_k)$ for $a \in \mathbb{C}$.

(9.8.18) Let $\mathcal{X}$ be a Banach space, and $\{x_k\} \subset \mathcal{X}$. Show that if $\sum_k x_k$ converges absolutely, then it converges unconditionally.

(9.8.19) Let $\mathcal{X}$ be a finite-dimensional Banach space, and $\{x_k\} \subset \mathcal{X}$. Show that if $\sum_k x_k$ converges unconditionally, then it converges absolutely.

(9.8.20) Let $p \in (1, \infty)$. Find a series in $\ell^p(\mathbb{N})$ that converges unconditionally but not absolutely.

## 9.9. A Brief Excursion into Injectivity

The contents of this section are certainly not needed anywhere else in the text, but they represent a nice blend of many of the ideas discussed so far. It is important to remark that the word "injective" has been used frequently, in the sense of "one-to-one"; the meaning here is different and unrelated, but both uses are standard terminology. In general there should be no confusion as the first use of the word refers to functions, while the new one refers to spaces.

Recall that a **linear isometry** between Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ is a linear map $\iota : \mathcal{X} \to \mathcal{Y}$ with $\|\iota(x)\| = \|x\|$ for all $x \in \mathcal{X}$. A very common example occurs when $\mathcal{X} \subset \mathcal{Y}$ and $\iota$ is the identity. A linear isometry can be seen as an embedding, for $\iota(\mathcal{X}) \simeq \mathcal{X}$ isometrically.

**9.9.1. Definition.**

*A Banach space $\mathcal{X}$ is said to be* ***injective*** *if for any Banach spaces $\mathcal{Y}$ and $\mathcal{Z}$ with a linear isometry $\iota : \mathcal{Y} \to \mathcal{Z}$ and $T \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$, there exists $\tilde{T} \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$ with $\|\tilde{T}\| = \|T\|$ and $\tilde{T} \circ \iota = T$.*

This is the usual diagram picturing the situation:

$$\begin{tikzcd} \mathcal{Z} \arrow[dr, dashed, "\tilde{T}"] & \\ \mathcal{Y} \arrow[u, "\iota"] \arrow[r, "T"'] & \mathcal{X} \end{tikzcd}$$

For a familiar example, the Hahn–Banach Theorem (Corollary 5.7.6) proves that $\mathbb{R}$ and $\mathbb{C}$ are injective as real and complex Banach spaces respectively.

The property we defined as injectivity is sometimes called 1-injectivity (also, **isometric injectivity**) in the literature. Those authors consider $\lambda$-injectivity where instead of preserving the norm, the extension is required to satisfy $\|\tilde{T}\| \leq$

$\lambda\,\|T\|$. In this cases, "injective" with no qualifier often poses no requirement on the norm of $\tilde T$. In this language, $c_0$ is 2-injective by Proposition 9.8.33.

It is a priori far from obvious which spaces are injective. But here are some results.

**9.9.2. Proposition.**

*Let $S$ be a set. Then $\ell^\infty(S)$ is injective.*

*Proof.* Let $\mathcal{Y},\mathcal{Z}$ be Banach spaces with a linear isometry $\iota:\mathcal{Y}\to\mathcal{Z}$. Let $T:\mathcal{Y}\to\ell^\infty(S)$ be bounded. For each $s\in S$, consider the map $T_s:\mathcal{Y}\to\mathbb{C}$ given by $T_sy=(Ty)_s$. As $T$ is bounded, we get that $T_s$ is a bounded linear functional. By Hahn–Banach (Corollary 5.7.6) there exists $\phi_s:\mathcal{Z}\to\mathbb{C}$, bounded, with $\|\phi_s\|=\|T_s\|$ and $\phi_s\circ\iota=T_s$. Define $\tilde T:\mathcal{Z}\to\ell^\infty(S)$ by $(\tilde Tz)(s)=\phi_s(z)$, $s\in S$. For $y\in\mathcal{Y}$ we have

$$\tilde T\circ\iota(y)=\{\phi_s(\iota(y))\}_s=\{T_sy\}_s=Ty,$$

so $\tilde T\circ\iota=T$. And, for $z\in\mathcal{Z}$ with $\|z\|\leq 1$,

$$\begin{aligned}\|\tilde Tz\|&=\sup\{|\phi_s(z)|:\ s\}\leq\sup\{\|\phi_s\|:\ s\}=\sup\{\|T_s\|:\ s\}\\&=\sup\{|(Ty)_s|:\ s,\ \ \|y\|=1\}\\&=\sup\{\|Ty\|:\ \|y\|=1\}\leq\|T\|.\end{aligned}$$

Then $\|\tilde T\|\geq\|T\|$; and as $\tilde T$ extends $T$ we get $\|\tilde T\|=\|T\|$. Thus $\ell^\infty(S)$ is injective. ☐

A consequence of Proposition 9.9.2 is that every Banach space embeds isometrically in an injective space. Indeed,

**9.9.3. Corollary.**

*Let $\mathcal{X}$ be a Banach space. Then there exists a set $S$ and a linear isometry $j:\mathcal{X}\to\ell^\infty(S)$. When $\mathcal{X}$ is separable, $S$ can be taken to be $\mathbb{N}$.*

*Proof.* Let $S=\overline{B_1^{\mathcal{X}^*}(0)}$ and $\mathcal{Y}=\ell^\infty(S)$. Define $\iota:\mathcal{X}\to\mathcal{Y}$ by $\iota(x)=\{\varphi(x)\}_{\varphi\in S}$. Then $\iota$ is linear, and isometric by Corollary 5.7.7.

When $\mathcal{X}$ is separable, let $\{x_n\}_n\subset\mathbb{N}$ be a countable dense subset. For each $n$, by Corollary 5.7.7 there exists $\varphi_n\in\mathcal{X}^*$ with $\|\varphi_n\|=1$ and $|\varphi_n(x_n)|=\|x_n\|$. Let $j:\mathcal{X}\to\ell^\infty(\mathbb{N})$ be given by $j(x)=(\varphi_n(x))_n$. This is trivially linear, and $\|j(x)\|\leq\|x\|$ for all $x$. For a fixed $x\in\mathcal{X}$ and any $\varepsilon>0$ there exists $n$ such that $\|x_n-x\|<\varepsilon$. Then

$$\|j(x)\|\geq|\varphi_n(x)|\geq|\varphi_n(x_n)|-|\varphi_n(x-x_n)|\geq\|x_n\|-\varepsilon\geq\|x\|-2\varepsilon.$$

As $\varepsilon$ was arbitrary, this shows the reverse inequality and so $\|j(x)\| = \|x\|$ for all $x \in \mathcal{X}$. Thus $j$ embeds $\mathcal{X}$ isometrically into $\ell^\infty(\mathbb{N})$. $\square$

From Corollary 9.9.3 we have that $\ell^\infty(\mathbb{N})$ is universal for separable Banach spaces. And from Theorem 7.6.6 we know that $C[0,1]$ is similarly universal; this is more interesting, in a sense, since $C[0,1]$ is separable. For a different kind of universality, we showed in Theorem 6.1.10 that $\ell^1(\mathbb{N})$ is universal in the sense that every separable Banach space is isometrically isomorphic to a quotient of it.

We will soon see that $L^\infty[0,1]$ is also injective. This next result works in more generality than stated (namely, for abelian $C^*$-algebras; see Propositions 13.2.22 and 13.2.24 and Theorem 13.2.19), but this is the context where we can prove it now.

***9.9.4. Lemma.***

*Let $S$ be a nonempty sets and $T : \ell^\infty(S) \to \ell^\infty(S)$ linear and such that $\|T\| = 1$ and $T(1) = 1$. Then $T$ is positive, meaning that if $g(s) \geq 0$ for all $s$, then $(Tg)(s) \geq 0$ for all $s$.*

*Proof.* We will prove first that $T$ maps real-valued functions to real-valued functions. Fix $g \in \ell^\infty(S)$ with $\|g\|_\infty = 1$ and $g(s) \in \mathbb{R}$ for all $s$. For any $n \in \mathbb{Z}$,

$$|(Tg)(s) + in| = \big|[T(g + in)](s)\big| \leq \|g + in\|_\infty = (1 + n^2)^{1/2}$$

This implies in particular that $n + \operatorname{Im}(Tg)(s) \leq (1+n^2)^{1/2}$ for all $n \in \mathbb{N}$. As $(1+n^2)^{1/2} - n \to 0$, it follows that $\operatorname{Im}(Tg)(s) \leq 0$. But the same estimates work for $-g$, and hence $(Tg)(s) \in \mathbb{R}$.

Now fix $g \in \ell^\infty(S)$ with $g(s) \geq 0$ for all $s$, and assume without loss of generality that $\|g\|_\infty = 1$. By the first part of the proof we know that $Tg$ is real. Since $0 \leq g \leq 1$, we have $-1 \leq 2g - 1 \leq 1$, so $\|2g - 1\|_\infty \leq 1$. Then $T(2g - 1) = 2(Tg) - 1$ is real-valued with $|2(Tg)(s) - 1| \leq 1$ for all $s$, which means that

$$-1 \leq 2(Tg)(s) - 1 \leq 1, \qquad s \in S.$$

This inequality is $0 \leq (Tg)(s) \leq 1$, which in particular implies that $Tg$ is positive. $\square$

Not surprisingly, linear isometries preserve injectivity:

**9.9.5. Lemma.**

*Let $\mathcal{X}$ be an injective Banach space, and $j : \mathcal{X} \to \mathcal{Y}$ a linear isometry. Then $j(\mathcal{X})$ is injective.*

*Proof.* Let $\mathcal{Y}, \mathcal{Z}$ be Banach spaces with a linear isometry $\iota : \mathcal{Y} \to \mathcal{Z}$ and $T \in \mathcal{B}(\mathcal{Y}, j(\mathcal{X}))$. Using the injectivity of $\mathcal{X}$ with the map $j^{-1} \circ T$, we obtain $\tilde{T} \in \mathcal{B}(\mathcal{Z}, \mathcal{X})$ with $\|\tilde{T}\| = \|j^{-1} \circ T\| = \|T\|$ and $\tilde{T} \circ \iota = j^{-1} \circ T$.

$$\begin{tikzcd}
\mathcal{Z} \arrow[rrd, dashed, "\tilde{T}"] & & \\
\mathcal{Y} \arrow[u, "\iota"] \arrow[r, "T"'] & j(\mathcal{X}) \arrow[r, "j^{-1}"'] & \mathcal{X}
\end{tikzcd}$$

Then $j \circ \tilde{T}$ is the desired extension, since $j \circ \tilde{T} \circ \iota = T$ and $\|j \circ \tilde{T}\| = \|\tilde{T}\| = \|T\|$. $\square$

Injectivity is certainly a useful property—after all, for several chapters now we have obtained a lot of mileage out of the fact that $\mathbb{C}$ is injective—but to have it available we first need to be able to show that some spaces are injective. A crucial tool for that is the following characterization.

**9.9.6. Proposition.**

*Let $\mathcal{X}$ be a Banach space. The following statements are equivalent:*

*(i) $\mathcal{X}$ is injective;*

*(ii) for any linear isometry $j : \mathcal{X} \to \mathcal{W}$ into a Banach space $\mathcal{W}$, the subspace $j(\mathcal{X}) \subset \mathcal{W}$ is complemented via a projection $P$ with $\|P\| = 1$;*

*(iii) there exists a linear isometry $j : \mathcal{X} \to \mathcal{W}$, where $\mathcal{W}$ is an injective Banach space, such that $j(\mathcal{X})$ is complemented.*

*Proof.* (i) $\implies$ (ii) Suppose that $\mathcal{X}$ is injective and that $j : \mathcal{X} \to \mathcal{W}$ is an isometry. By Lemma 9.9.5 we have that $j(\mathcal{X})$ is injective. By this injectivity we can extend the identity $\iota : j(\mathcal{X}) \to j(\mathcal{X})$ to a bounded map $P : \mathcal{W} \to j(\mathcal{X})$ with $\|P\| = \|\iota\| = 1$ and $P \circ j = j$. We have $P^2 = P$, so $P$ is a norm-one projection onto $j(\mathcal{X})$.

(ii) $\implies$ (iii) Take $\mathcal{W} = \mathcal{X}$.

(iii) $\implies$ (i)We have a norm-one projection $P$ on $\mathcal{W}$ with $P\mathcal{W} = j(\mathcal{X})$. We want to show that $\mathcal{X}$ is injective. Let $\mathcal{Y}, \mathcal{Z}$ with isometry $\iota : \mathcal{Y} \to \mathcal{Z}$ and $T \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$. As $\mathcal{W}$ is injective and $j \circ T$ maps $\mathcal{Y}$ into $\mathcal{W}$, there exists

$\hat{T} : \mathcal{Z} \to \mathcal{W}$ with $\hat{T} \circ \iota = j \circ T$ and $\|\hat{T}\| = \|j \circ T\| = \|T\|$. Let $\tilde{T} = j^{-1} \circ P \circ \hat{T}$. Then

$$\tilde{T} \circ \iota = j^{-1} \circ P \circ \hat{T} \circ \iota = j^{-1} \circ P \circ j \circ T = j^{-1} \circ j \circ T = T,$$

and

$$\|\tilde{T}\| = \|j^{-1} \circ P \circ \hat{T}\| = \|P \circ \hat{T}\| \leq \|\hat{T}\| = \|T\|.$$

We also have, for $x \in \mathcal{X}$, that $\tilde{T}(\iota(x)) = Tx$, so $\|\tilde{T}\| \geq \|T\|$, showing that $\|\tilde{T}\| = \|T\|$. Thus $\mathcal{X}$ is injective. $\square$

As $c_0$ is not complemented in $\ell^\infty(\mathbb{N})$ (Proposition 5.3.6), Proposition 9.9.6 shows that $c_0$ is not injective, even though it is 2-injective by Proposition 9.8.33. More generally, no separable infinite-dimensional Banach space is injective (we will show this in Corollary 9.9.15). In particular $\ell^p(\mathbb{N})$ is not injective for any $p \in [1, \infty)$.

To somehow complement Proposition 9.9.2,

***9.9.7. Proposition.***

*$L^\infty[0,1]$ is injective.*

*Proof.* Recall that every $f \in L^\infty[0,1]$ is a uniform limit of simple functions (Theorem 2.4.13). So if we construct the subspaces

$$\mathcal{W}_{E_1,\dots,E_m} = \operatorname{span}\{1_{E_1}, \dots, 1_{E_m}\} \subset L^\infty[0,1],$$

where $E_1, \dots, E_m$ run over all measurable partitions of $[0,1]$, ordered by inclusion, we obtain an increasing net such that the union is dense. Let $j : L^\infty[0,1] \to \mathcal{X}$ be a linear isometry; we want to show that $j(L^\infty[0,1])$ is complemented. Each $\mathcal{W} = \mathcal{W}_{E_1,\dots,E_m}$ is isomorphic to $\ell^\infty(\{1,\dots,m\})$ and so injective by Proposition 9.9.2. By Proposition 9.9.6 there exists a projection $P_{\mathcal{W}} : \mathcal{X} \to j(\mathcal{W})$ with $\|P_{\mathcal{W}}\| = 1$. Every $P_{\mathcal{W}}$ is in the unit ball of $\mathcal{B}(\mathcal{X}, j(L^\infty[0,1]))$, and this unit ball is pointwise weak*-compact by Proposition 9.4.10. Thus the net $\{P_{\mathcal{W}}\}$ admits a pointwise weak*-convergent subnet, convergent to some $P \in \mathcal{B}(\mathcal{X}, j(L^\infty[0,1]))$. Since each $P_{\mathcal{W}}$ is the identity on $j(\mathcal{W})$ and the union of these is dense in $j(L^\infty[0,1])$, we get that $P$ is the identity on $j(L^\infty[0,1])$. Thus $L^\infty[0,1]$ is injective by Proposition 9.9.6. $\square$

That finite subspaces are complemented was done in Exercise 5.7.2. But there was no obvious way there to control the norm of the projection. It is also worth mentioning that the argument in Proposition 9.9.7 does not use $[0,1]$ in a very particular way. In fact, the argument works verbatim for $L^\infty(X)$ for any $\sigma$-finite measure space.

So far the only injective spaces we have identified are $\ell^\infty(S)$ and $L^\infty(X)$ for a $\sigma$-finite measure space $X$; and it is not clear at all which other spaces are injective and which are not. Luckily work of Nachbin [Nac49], Goodner [Goo50], and Kelley [Kel52] characterized injectivity for real Banach spaces, and Hasumi [Has58] and later Cohen [Coh64] addressed the complex case. The basic idea is to repeat the technique that works to prove the Hahn–Banach theorem, now in a “Banach-space-valued” version. Several of the technicalities of this approach were worked in Lemma 7.4.29, Proposition 7.4.30, and Theorem 7.4.31.

Meanwhile, for spaces of continuous functions we have a complete characterization of injectivity:

**9.9.8. *Theorem.***

*Let $X$ be a compact Hausdorff space. The following statements are equivalent:*

*(i) $C(X)$ is injective;*

*(ii) $C(X)$ is order-complete;*

*(iii) $X$ is extremally disconnected.*

*Proof.* (i) $\Longrightarrow$ (ii) Since $C(X) \subset \ell^\infty(X)$ isometrically, by the injectivity of $C(X)$ and Proposition 9.9.6 there is a norm-one projection $P : \ell^\infty(X) \to C(X)$. By Lemma 9.9.4, this projection preserves positivity. Let $\mathcal{R} \subset C_{\mathbb{R}}(X)$ be bounded. Let $g \in \ell^\infty(X)$ be given by $g(x) = \sup\{f(x) : \ f \in \mathbb{R}\}$. The fact that $\mathcal{R}$ is bounded gives $g \in \ell^\infty(X)$. Let $h = Pg$. Because $P$ preserves positivity, if $f \in \mathcal{R}$ then $f \leq g$, which means that $g - f \geq 0$; then $h - f = P(g - f) \geq 0$ and so $h$ is an upper bound for $\mathcal{R}$ in $C_{\mathbb{R}}(X)$. And if $k \in C_{\mathbb{R}}(X)$ is an upper bound for $\mathcal{R}$, we have that $k$ is also an upper bound for $\mathcal{R}$ in $\ell^\infty(X)$. Thus $g \leq k$ and $k - h = P(k - g) \geq 0$, so $h$ is the least upper bound for $\mathcal{R}$ in $C_{\mathbb{R}}(X)$.

(ii) $\Longleftrightarrow$ (iii) This is Proposition 7.4.27.

(iii) $\Longrightarrow$ (i) Use Theorem 7.4.31 with the seminorm $q(f) = \|f\|_\infty$. $\square$

We know from Proposition 9.9.2 and Corollary 9.9.3 that every Banach space $\mathcal{X}$ is contained in an injective Banach space. It then makes sense to consider the smallest injective space that contains $\mathcal{X}$.

**9.9.9. Definition.**

*Let $\mathcal{X}$ be a Banach space. An* ***injective envelope*** *for $\mathcal{X}$ is a pair $(\iota, I(\mathcal{X}))$ where $I(\mathcal{X})$ is an injective Banach space and $\iota : \mathcal{X} \to I(\mathcal{X})$ is a linear isometry, with the property that no proper subspace of $I(\mathcal{X})$ that contains $\iota(\mathcal{X})$ is injective.*

The existence of the injective envelope for Banach spaces was established by Cohen [Coh64], with the techniques that appear in the proof of Theorem 9.9.14. The arguments we present for the existence, though, follow ideas of Kaufman and Hamana [Kau66, Ham79a, Ham79b]. It is interesting to remark that, as opposed to many cases of "object of this type generated by a set" (vector spaces, groups, algebras, $\sigma$-algebras, etc., etc.), the intersection of injective spaces is not necessarily injective (Exercise 9.9.1). This is what makes a non-trivial argument necessary to show that injective envelopes exist.

**9.9.10. Theorem. (Existence of the Injective Envelope)**

*Let $\mathcal{X}$ be a Banach space. Then there exists an injective envelope $(\iota, I(\mathcal{X}))$ for $\mathcal{X}$.*

*Proof.* Let $S$ and $j$ as in Corollary 9.9.3, and put $\mathcal{Y} = \ell^\infty(S)$. Given a linear contractive map $\psi : \mathcal{Y} \to \mathcal{Y}$, we associate to it a seminorm $p_\psi$ given by $p_\psi(y) = \|\psi(y)\|$. Consider the family

$$\mathcal{F} = \{p_\phi : \ \phi : \mathcal{Y} \to \mathcal{Y} \ \text{ linear, contractive, } \phi|_{\iota(\mathcal{X})} = \mathrm{id}_{\iota(\mathcal{X})}\},$$

with the partial order $p_{\phi_1} \leq p_{\phi_2}$ if $\|\phi_1(y)\| \leq \|\phi_2(y)\|$ for all $y$. We want to use Zorn's Lemma to show that $\mathcal{F}$ admits a minimal element. Consider a non-increasing chain $\{p_{\psi_j}\} \subset \mathcal{F}$. The maps $\psi_j$ belong in $\mathcal{B}(\mathcal{Y}) = \mathcal{B}(\mathcal{Y}, \mathcal{Y}_0^*)$, where $\mathcal{Y}_0 = \ell^1(S)$ (Exercise 5.6.6). By Proposition 9.4.10, the chain admits a weak*-pointwise cluster point $\psi \in \mathcal{B}(\mathcal{Y})$. That is, there exists a subnet $\{\psi_{j_k}\}$ such that for all $y \in \mathcal{Y}$ and $b \in \mathcal{Y}_0$, $\langle \psi_{j_k}(y), b\rangle \to \langle \psi(y), b\rangle$. We have, by definition of the dual norm,

$$\|\psi(y)\| = \sup\{|\langle \psi(y), b\rangle| : \ b \in \mathcal{Y}_0, \ \|b\| = 1\}.$$

Given $\varepsilon > 0$, $b \in \mathcal{Y}_0$ with $\|b\| = 1$, and some index $j_{k'}$, there exists $k'' > k'$ such that $|\langle \psi(y) - \psi_{j_k}(y), b\rangle| < \varepsilon$ for all $k > k''$. Then

$$|\langle \psi(y), b\rangle| \leq \varepsilon + |\langle \psi_{j_k}(y), b\rangle| \leq \varepsilon + \|\psi_{j_k}(y)\| \leq \varepsilon + \|\psi_{j_{k'}}(y)\|.$$

As $\varepsilon$ and $b$ were arbitrary, it follows that $\|\psi(y)\| \leq \|\psi_{j'_k}(y)\|$ for all $y$, and so $p_\psi \leq p_{\psi_{j_k}}$ for all $k$. As the chain $\{p_{\psi_j}\}$ is totally ordered, $p_\psi \leq p_{\psi_j}$ for all $j$ and hence $p_\psi$ is a lower bound for the chain. Then Zorn's Lemma guarantees that $\mathcal{F}$ admits a minimal element $p_\phi$.

We will show that $\phi$ is a projection. Since $\phi$ is contractive, $\|\phi^2(y)\| \leq \|\phi(y)\|$ for all $y \in \mathcal{Y}$. So $p_{\phi^2} \leq p_\phi$, but $p_\phi$ is minimal so $p_{\phi^2} = p_\phi$. That is, $\|\phi^2(y)\| = \|\phi(y)\|$ for all $y$, and iterating this equality we get $\|\phi^k(y)\| = \|\phi(y)\|$. Consider the map $\phi_n = \frac{1}{n}(\phi + \phi^2 + \cdots + \phi^n)$. The triangle inequality and the equalities $\|\phi^k(y)\| = \|\phi(y)\|$ give us $\|\phi_n(y)\| \leq \|\phi(y)\|$. The minimality of $p_\phi$ then implies that $\|\phi_n(y)\| = \|\phi(y)\|$ for all $y$. Now

$$\|\phi^2(y) - \phi(y)\| = \|\phi(\phi(y) - y)\| = \|\phi_n(\phi(y) - y)\|$$

$$= \frac{1}{n} \|\phi^{n+1}(y) - \phi(y)\| \leq \frac{2\|y\|}{n}.$$

As this can be done for all $n$, $\phi^2(y) = \phi(y)$ for all $y \in \mathcal{Y}$. Then $\phi$ is a norm-1 projection; by Proposition 9.9.6, $\phi(\mathcal{Y})$ is an injective Banach space that contains $\iota(\mathcal{X})$.

Now suppose that $\iota(\mathcal{X}) \subset \mathcal{Z} \subset \phi(\mathcal{Y})$ for some injective subspace $\mathcal{Z}$. By Proposition 9.9.6 there exists a contractive projection $\psi : \mathcal{Y} \to \mathcal{Z}$. Then $\|\psi(\phi(y))\| \leq \|\phi(y)\|$ for all $y$, so by the minimality of $\phi$ we have $\|\psi(\phi(y))\| = \|\phi(y)\|$ for all $y$. Let $z' \in \phi(\mathcal{Y}) \setminus \mathcal{Z}$ and put $z = (\phi - \psi)(z')$. We have $z \neq 0$, for otherwise $z' = \phi(z') = \psi(z') \in \mathcal{Z}$. Then

$$\|z\| = \|\phi(z)\| = \|\psi(\phi(z))\| = \|\psi(z)\| = 0$$

So $z = 0 \in \mathcal{Z}$, a contradiction. The conclusion is that $\mathcal{Z} = \phi(\mathcal{Y})$ and $(\iota, \phi(\mathcal{Y}))$ is an injective envelope for $\mathcal{X}$. □

The reason one has to go through the seminorms in the proof of Theorem 9.9.10 instead of applying Zorn directly to the maps $\psi$ is that it is not clear how to get a partial order on them. Concretely, it is natural to order them via their associated seminorms, but when we see it as an order of the operators, what one gets is not antisymmetric and thus not an order.

A crucial property of the injective envelope is its **rigidity**:

**9.9.11. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $j : \mathcal{X} \to \mathcal{Y}$ a linear isometry. The triple $(\mathcal{X}, j, \mathcal{Y})$ is **rigid** if the only contractive linear map $\psi : \mathcal{Y} \to \mathcal{Y}$ such that $\psi \circ j = j$ is $\mathrm{id}_{\mathcal{Y}}$.*

***9.9.12. Proposition. (Rigidity)***

*Let $\mathcal{X}, \mathcal{J}$ be Banach spaces and $j : \mathcal{X} \to \mathcal{J}$ a linear isometry. The following statements are equivalent:*

*(i) $(j, \mathcal{J})$ is an injective envelope for $\mathcal{X}$;*

*(ii) $(\mathcal{X}, j, \mathcal{J})$ is rigid.*

*In particular, the injective envelope is unique up to isometric isomorphism. And if $g : I(\mathcal{X}) \to \mathcal{K}$ is an isometric isomorphism of Banach spaces, then $(g \circ \iota, \mathcal{K})$ is an injective envelope for $\mathcal{X}$.*

*Proof.* (i) $\Longrightarrow$ (ii) Let $\psi : \mathcal{J} \to \mathcal{J}$ linear, contractive, and with $\psi \circ j = j$. Apply Corollary 9.9.3 to $\mathcal{J}$ to get a set $S$ and an isometric embedding $\alpha : \mathcal{J} \to \mathcal{Y}$, where $\mathcal{Y} = \ell^\infty(S)$. Since $\mathcal{J}$ is injective there exists a minimal seminorm $p_\alpha$, given by a projection $\phi_\alpha$, such that $\mathcal{J} = \phi_\alpha(\mathcal{Y})$ (the proof of Theorem 9.9.10 gives us $\alpha(\mathcal{J}) \subset \phi_\alpha(\mathcal{Y})$, but since $\alpha(\mathcal{J})$ is injective the inclusion has to be an equality for otherwise $\phi_\alpha(\mathcal{Y})$ would not be an injective envelope for $\mathcal{J}$). Since $\|\psi(\phi_\alpha(y))\| \leq \|\phi_\alpha(y)\|$ for all $y$ and $p_\alpha$ is minimal, $\|\psi(\phi_\alpha(y))\| = \|\phi_\alpha(y)\|$ for all $y$. Then $\ker \phi_\alpha = \ker \psi \circ \phi_\alpha$. And since $\operatorname{ran} \psi \circ \phi_\alpha \subset \operatorname{ran} \phi_\alpha$, we have $\psi \circ \phi_\alpha = \phi_\alpha$ by Exercise 9.6.16. That is, $\psi = \mathrm{id}_{\mathcal{J}}$.

(ii) $\implies$ (i) Suppose that $(\mathcal{X}, j, \mathcal{J})$ is rigid. Let $\mathcal{Z} \subset \mathcal{J}$ be an injective Banach subspace with $j(\mathcal{X}) \subset \mathcal{Z}$. Let $\psi = \mathrm{id}_{\mathcal{Z}}$ an extend it, by the injectivity of $\mathcal{Z}$, to a map $\tilde{\psi} : \mathcal{J} \to \mathcal{Z}$. We have $\tilde{\psi} \circ j = j$, since $j(\mathcal{X}) \subset \mathcal{Z}$. Then the rigidity implies that $\tilde{\psi} = \mathrm{id}_{\mathcal{Z}}$, and therefore $\mathcal{Z} = \mathcal{J}$. This shows that $(j, \mathcal{J})$ is an injective envelope for $\mathcal{X}$.

Now for the uniqueness, if $(k, \mathcal{K})$ is another injective envelope the injectivity of $\mathcal{J}$, applied to the map $j : \mathcal{X} \to \mathcal{J}$ and the linear isometry $k : \mathcal{X} \to \mathcal{K}$, gives us a linear contraction $\rho : \mathcal{K} \to \mathcal{J}$ with $\rho \circ k = j$. Similarly, there exists a linear contraction $\nu : \mathcal{J} \to \mathcal{K}$ with $\nu \circ j = k$. Then

$$(\rho \circ \nu) \circ j = \rho \circ (\nu \circ j) = \rho \circ k = j.$$

The rigidity of $(j, \mathcal{J})$ then implies that $\rho \circ \nu = \mathrm{id}_{\mathcal{J}}$. Similarly, $\nu \circ \rho = \mathrm{id}_{\mathcal{K}}$. As $\rho$ and $\nu$ are contractions,

$$\|y\| = \|\rho(\nu(y))\| \leq \|\nu(y)\| \leq \|y\|,$$

and so $\nu$ is an isometry. With an analog argument, $\rho$ is an isometry too.

We leave the last assertion, when $\mathcal{K}$ is isometrically isomorphic to $I(\mathcal{X})$, as an exercise (Exercise 9.9.2). $\square$

With the idea that the injective envelope of $\mathcal{X}$ is the smallest injective Banach space that contains $\mathcal{X}$ it should be possible, every time that $\mathcal{X}$ embeds in an injective space, to put an injective envelope in between. And that's the case indeed:

**9.9.13. Corollary.**

*Let $\mathcal{X}$ be a Banach space and $\mathcal{Y}$ an injective Banach space with $\mathcal{X} \subset \mathcal{Y}$. Then there exists an injective subspace $\mathcal{Z} \subset \mathcal{Y}$ such that $(\mathrm{id}, \mathcal{Z})$ is an injective envelope for $\mathcal{X}$.*

*Proof.* Using the injectivity of $I(\mathcal{X})$ and $\mathcal{Y}$ respectively,

$$\begin{tikzcd} \mathcal{Y} \arrow[dr, dashed, "h"] & \\ \mathcal{X} \arrow[u, "\mathrm{id}"] \arrow[r, "j"'] & I(\mathcal{X}) \end{tikzcd} \qquad \begin{tikzcd} I(\mathcal{X}) \arrow[dr, dashed, "g"] & \\ \mathcal{X} \arrow[u, "j"] \arrow[r, "\mathrm{id}"'] & \mathcal{Y} \end{tikzcd}$$

there exist linear contractions $h : \mathcal{Y} \to I(\mathcal{X})$ and $g : I(\mathcal{X}) \to \mathcal{Y}$ such that $h|_{\mathcal{X}} = j$ and $g \circ j = \mathrm{id}_{\mathcal{X}}$. Then $h \circ g : I(\mathcal{X}) \to I(\mathcal{X})$ is a linear contraction that satisfies $(h \circ g) \circ j = h \circ \mathrm{id}_{\mathcal{X}} = j$. By the rigidity of the injective envelope we have $h \circ g = \mathrm{id}_{I(\mathcal{X})}$. In particular

$$\|x\| = \|h(g(x))\| \leq \|g(x)\| \leq \|x\|, \qquad x \in I(\mathcal{X}).$$

So $g$ is an isometry. By Exercise 9.9.2, $\mathcal{Z} = g(I(\mathcal{X}))$ makes $(g \circ j, \mathcal{Z})$ an injective envelope for $\mathcal{X}$, with $\mathcal{Z} \subset \mathcal{Y}$ and $g \circ j = \mathrm{id}_{\mathcal{X}}$. □

Theorem 9.9.14 was proven by Kelley [**Kel52**] in the real case. Extensions to the complex case were proven by Hasumi [**Has58**] and Cohen [**Coh64**]. We will follow the latter. Cohen's proof of Kelley's Theorem below will amount to finding a presentation of the injective envelope of $\mathcal{X}$ as $C(M)$, and this in turn is achieved via Gleason's map as in Theorem 7.4.36.

***9.9.14. Theorem. (Kelley/Cohen)***

*Let $\mathcal{X}$ be a Banach space. The following statements are equivalent:*

*(i) $\mathcal{X}$ is injective;*

*(ii) $\mathcal{X}$ is isometrically isomorphic to $C(K)$ with $K$ extremally disconnected.*

*Proof.* (i) $\implies$ (ii) We already had a situation where we needed to find a topological space out of nowhere in Theorem 7.4.14. Here again we will look into the dual. The weak*-topology automatically makes each $x \in \mathcal{X}$ a continuous function on the (compact!) closed unit ball of $\mathcal{X}^*$; then we need to thin down the compact set so that no function not coming from $\mathcal{X}$ arises; this will require the injectivity. Let $K_0 = \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$, considered with the weak*-topology. It is clear that $K_0$ is balanced. We will say that a subset $R \subset \mathcal{X}^*$ is **circled** if $\mathbb{T}R \subset R$ and **removed** if for each $r \in R$ there exists $\lambda \in \mathbb{T}$ such that $(\mathbb{T} \setminus \{\lambda\})r \in R$ but $\lambda r \notin R$.

Let $D = \mathbb{D} \setminus [0,1)$. Fix $W \subset K_0$ nonempty and circled, and $\varphi \in W$ and $x \in \mathcal{X}$ with $\varphi(x) \in D$. Then $W' = (\hat{x})^{-1}(D) \cap W$ is nonempty (as it contains $\varphi$), open, and removed (because if $\varphi(x) = e^{i\theta}|\varphi(x)|$, then $e^{-i\theta}\varphi \notin W'$). Now we can consider the family of all nonempty open removed subsets of $K_0$, and use Zorn's Lemma to get $V$ maximal in $K_0$ with respect to the property of being nonempty, open, and removed. This works because if $\{V_j\}$ is a monotone increasing family of nonempty open removed sets, we can take $V' = \bigcup_j V_j$. Given $v \in V'$ there exists $j$ such that $v \in V_j$; this means that there exists a unique $\lambda_j \in \mathbb{T}$ with $\lambda_j v \notin V_j$; if we had $\lambda_j v \in V_k$ for $k > j$ this would mean that there is a different scalar $\mu \in \mathbb{T}$ with $\mu v \notin V_k$, but this would imply that $\mu v \notin V_j$. It follows that for each $v \in V$ there is a unique

$\lambda \in \mathbb{T}$ with $\lambda v \notin V$; that is, $V$ is removed. So $V$ is an upper bound for the chain and then Zorn's Lemma guarantees that there is a maximal nonempty open removed subset $V$ of $K_0$.

It follows that $\mathbb{T}V$ is open and dense in $K_0$. Indeed, if $\mathbb{T}V$ is not dense, there exists $W \subset K_0$ with $W \cap \mathbb{T}V = \varnothing$. As in the previous paragraph, there exists $W' \subset W$ nonempty, open, and removed, which contradicts the maximality of $V$. So $\mathbb{T}V$ is open and dense, and then $\mathbb{T}V \cap \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$ is dense in $\operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$. Let

$$U = \{\varphi \in \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)} : \ \varphi \notin V, \ (\mathbb{T} \setminus \{1\})\varphi \subset V\},$$

and let $K = \overline{U}^{w^*}$. The fact that $U \neq \varnothing$ follows from the density of $\mathbb{T}V \cap \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$ and that $V$ is removed. We have

(1) $K$ is weak*-compact, and $U \subset \mathcal{X} \cap \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$;
(2) $\overline{\mathbb{T}U}^{w^*} = \overline{\operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}}^{w^*}$;
(3) for each $\varphi \in U$, $K \cap \mathbb{T}\{\varphi\} = \{\varphi\}$.

The first property is satisfied by construction. Property (2) again follows from the density of $\mathbb{T}V \cap \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$ in $\operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$. And (3) is obtained from $U \cap V = \varnothing$, which implies $K \cap V = \varnothing$. The only new element that appears when multiplying by all unit scalars is $\varphi$.

By Gleason's Theorem (7.4.36) there exists a minimal Stonean space $M$ with a continuous surjection $\gamma : M \to K$. Let $j : \mathcal{X} \to C(M)$ be given by $j(x)(m) = \gamma(m)(x)$ (recall that $K \subset \mathcal{X}^*$). The continuity of $\gamma$, which is on the weak*-topology on its codomain, guarantees that $j$ is continuous. The linearity of $j$ is straightforward from the fact that $\gamma(m) \in \mathcal{X}^*$. Given $x \in \mathcal{X}$, choose $\psi \in \operatorname{Ext} \overline{B_1^{\mathcal{X}^*}(0)}$ with $|\psi(x)| = \|x\|$. By (2) there exists a net $\{\varphi_j\} \subset \mathbb{T}U$ with $\varphi_j \xrightarrow{\text{weak}^*} \psi$. In particular, $\|x\| = \lim_j |\varphi_j(x)|$. Fix $\varepsilon > 0$ and choose an index $k$ with $\|x\| < |\varphi_k(x)| + \varepsilon$. Let $\lambda \in \mathbb{T}$ with $\lambda\varphi_k \in U$. As $\lambda\varphi_k \in \mathcal{X}$, there exists $m \in M$ with $\gamma(m) = \lambda\varphi_k$. Then

$$\|x\| - \varepsilon < |\varphi_k(x)| = |\lambda\varphi_k(x)| = |\gamma(m)(x)| = |j(x)(m)| \leq \|j(x)\|.$$

As this can be done for all $\varepsilon > 0$, we get $\|j(x)\| \geq \|x\|$. The reverse inequality $\|j(x)\| \leq \|x\|$ comes directly from the fact that $\gamma(m) \in K \subset \overline{B_1^{\mathcal{X}^*}(0)}$. Hence $j$ is an isometry. We know from Theorem 9.9.8 that $C(M)$ is injective.

Because $j(\mathcal{X})$ is injective, there exists a contractive projection map $P : C(M) \to C(M)$ with $P|_{j(\mathcal{X})} = \mathrm{id}_{j(\mathcal{X})}$. If we prove that $P$ is an isometry, then $j(\mathcal{X}) = C(M)$, for $\ker P = \{0\}$ and then $P = \mathrm{id}_{C(M)}$. So we will be complete the proof by showing that $P$ is an isometry.

Fix $f \in C(M)$ and $\varepsilon > 0$. Assume without loss of generality that $\|f\| \in f(M)$ (we can otherwise replace $f$ with a scalar multiple of it, without changing the norm). Let $M_\varepsilon = f^{-1}(B_\varepsilon(\|f\|))$, open. As $\gamma$ is minimal, $\gamma(M \setminus M_\varepsilon) \subsetneq K$; and since $U$ is dense in $K$, there exists $\varphi \in U \cap (K \setminus \gamma(M \setminus$

$M_\varepsilon)) = U \cap \gamma(M_\varepsilon)$. Define $\nu : M \to C(M)^*$ by $\nu(m) = \delta_m$. This map is continuous (since we use the weak*-topology on $C(M)^*$) and injective (since $C(M)$ separates points on $M$), so a homeomorphism onto its image. The set $\nu(\gamma^{-1}(\{\varphi\}))$ consists of linear functionals $\psi$ such that $|\psi(f) - \|f\|| < \varepsilon$. Hence

$$\text{if } \psi \in \overline{\operatorname{conv}}^{w^*} \nu(\gamma^{-1}(\{\varphi\})), \qquad \text{then } |\psi(f)| \geq \|f\| - \varepsilon. \tag{9.48}$$

Given $\psi \in (j^*)^{-1}(\{\varphi\}) \cap \operatorname{Ext} \overline{B_1^{C(M)^*}(0)}$, by Exercise 7.5.16 we know that $\psi = \lambda\delta_m$ for some $\lambda \in \mathbb{T}$ and $m \in M$. Therefore, using that $\delta_m \circ j = \gamma(m)$ by definition,

$$\varphi = j^*(\psi) = \lambda\, j^* \circ \delta_m = \lambda\, \delta_m \circ j = \lambda\, \gamma(m).$$

By property (3), we get that $\lambda = 1$ and $\gamma(m) = \varphi$. Then $m \in \gamma^{-1}(\{\varphi\})$ and $\psi = \nu(m) \in \nu(\gamma^{-1}(\{\varphi\}))$. Thus

$$\overline{\operatorname{conv}}^{w^*} \big((j^*)^{-1}(\{\varphi\}) \cap \operatorname{Ext} \overline{B_1^{C(M)^*}(0)}\big) \subset \overline{\operatorname{conv}}^{w^*} \nu(\gamma^{-1}(\{\varphi\})). \tag{9.49}$$

We can extend the bounded linear functional $\varphi \circ j^{-1} : j(\mathcal{X}) \to \mathbb{C}$ to $C(M)$ via Hahn–Banach (Corollary 5.7.6) to $\tilde{\varphi} : C(M) \to \mathbb{C}$. Then

$$j^*(P^*\tilde{\varphi}) = j^*(\tilde{\varphi} \circ P) = \tilde{\varphi} \circ P \circ j = \tilde{\varphi} \circ j = \varphi.$$

This means that, using Krein–Milman (Theorem 7.5.11),

$$P^*\tilde{\varphi} \in (j^*)^{-1}(\{\varphi\}) \cap \overline{B_1^{C(M)^*}(0)} \subset \overline{\operatorname{conv}}^{w^*} \big((j^*)^{-1}(\{\varphi\}) \cap \operatorname{Ext} \overline{B_1^{C(M)^*}(0)}\big).$$

Combining (9.49) and (9.48),

$$\|f\| - \varepsilon \leq |(P^*\tilde{\varphi})(f)| = |\tilde{\varphi}(Pf)| \leq \|Pf\| \leq \|f\|.$$

As this whole convoluted process can be repeated for any $\varepsilon > 0$, we have $\|Pf\| = \|f\|$.

(ii) $\implies$ (i) This is Theorem 9.9.8. ☐

A consequence of Theorem 9.9.14 and Proposition 9.9.7 is that there exists a compact Hausdorff space $X$ such that $L^\infty[0,1] \simeq C(X)$. This isomorphism is an isometry of Banach spaces. But we will see in Theorem 11.2.3, combined with Gelfand–Kolmogorov (Theorem 7.4.8) that there is actually an isometric algebra isomorphism between $L^\infty[0,1]$ and $C(X)$.

We can now show that many common Banach spaces are not injective.

**9.9.15. Corollary.**

*An injective separable space is isometrically isomorphic to $\ell^\infty(S)$ with $S$ finite. In particular, no separable infinite-dimensional Banach space is injective.*

*Proof.* Let $\mathcal{X}$ be separable and injective. By Theorem 9.9.14 there exists a Stonean space $M$ such that $\mathcal{X}$ is isometrically isomorphic to $C(M)$. As $\mathcal{X}$ is separable so is $C(M)$. By Proposition 7.4.24, $M$ is metrizable, and then by Proposition 7.4.28, $M$ is finite. So $C(M) = \ell^\infty(M)$, with $M$ is finite. $\square$

We immediately get that none of the usual Banach spaces like $c_0$ and $\ell^p(\mathbb{N})$ with $1 \leq p < \infty$ are injective. Even more strikingly, by Corollary 5.2.6 no $\ell^p(\{1, \ldots, n\})$ with $p \in [1, \infty)$ is injective.

Here is another application of Pełczyński's decomposition.

**9.9.16. Theorem.**

*As Banach spaces, $\ell^\infty(\mathbb{N}) \simeq L^\infty[0,1]$. That is, there exist bounded linear maps $T : \ell^\infty(\mathbb{N}) \to L^\infty[0,1]$ and $S : L^\infty[0,1] \to \ell^\infty(\mathbb{N})$ with $TS = \mathrm{id}_{L^\infty[0,1]}$ and $ST = \mathrm{id}_{\ell^\infty(\mathbb{N})}$.*

*Proof.* Let $\mathcal{X} = \ell^\infty(\mathbb{N})$, $\mathcal{Y} = L^\infty[0,1]$. Define $i : \mathcal{X} \to \mathcal{Y}$ by

$$i(x) = \sum_{n=1}^\infty x_n \, 1_{\left[\frac{1}{2n+1}, \frac{1}{2n}\right]}.$$

The series above should not be seen as a limit, but rather as shorthand notation for defining a piecewise function. The function $i$ is a linear isometry (Exercise 9.9.3). It is clear that it is not surjective, as no continuous function lies on its image. As $\ell^\infty(\mathbb{N})$ is injective (Proposition 9.9.2), the image of $i$ is complemented in $\mathcal{Y}$ by Proposition 9.9.6.

Let $\{g_n\} \subset L^1[0,1]$ be dense in its unit sphere. Define $j : \mathcal{Y} \to \mathcal{X}$ by

$$j(f) = \left( \int_{[0,1]} g_n \, f \right)_n.$$

Because $\{g_n\}$ is dense in the unit sphere of $L^1[0,1]$, $\|j(f)\| = \|f\|_\infty$ (via Proposition 5.6.8). As $j$ is also linear, it gives us an isometric embedding $\mathcal{Y} \to \mathcal{X}$. Since $\mathcal{Y}$ is injective (Proposition 9.9.7), $j(\mathcal{Y})$ is complemented in $\mathcal{X}$ by Proposition 9.9.6.

Since

$$\mathcal{X} \simeq \ell^\infty(2\mathbb{N}) \oplus \ell^\infty(1 + 2\mathbb{N}) \simeq \ell^\infty(\mathbb{N}) \oplus \ell^\infty(\mathbb{N}) = \mathcal{X} \oplus \mathcal{X}$$

and

$$\mathcal{Y} = L^\infty[0,1/2] \oplus L^\infty[1/2,1] \simeq L^\infty[0,1] \oplus L^\infty[0,1] = \mathcal{Y} \oplus \mathcal{Y},$$

we conclude by Pełczyński's decomposition (Proposition 9.8.22) that $\mathcal{X} \simeq \mathcal{Y}$. $\square$

*9.9.17. Remark.* While we showed in Theorem 9.9.16 that $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$ are isomorphic as Banach spaces, there is no isometric isomorphism between them. We will prove this in Corollary 11.2.12 when more machinery is available.

### 9.9.1. Exercises.

(9.9.1) Let $\mathcal{X}$ be a Banach space and $\mathcal{Y} \subset \mathcal{X}$ a closed subspace that is not complemented. Let $\mathcal{Z} = (\mathcal{X} \oplus_1 \mathcal{X})/\mathcal{K}$, where $\mathcal{K} = \{(y, -y) : \ y \in \mathcal{Y}\}$. Show that $\mathcal{X} \oplus 0$ and $0 \oplus \mathcal{X}$ are complemented in $\mathcal{Z}$ but their intersection is not.

(9.9.2) Let $\mathcal{X}$ be a Banach space, $(j, \mathcal{J})$ an injective envelope for $\mathcal{X}$, and $\mathcal{K}$ a Banach space with $g : \mathcal{J} \to \mathcal{K}$ an isometric isomorphism. Show that $(g \circ j, \mathcal{K})$ is an injective envelope for $\mathcal{X}$.

(9.9.3) Show that $i$, as in the proof of Theorem 9.9.16, is a linear isometry.

CHAPTER

# 10

# Bounded operators on a Hilbert space: Part I

We will focus on $\mathcal{B}(\mathcal{H})$, the space of linear bounded operators on a Hilbert space $\mathcal{H}$. We know from Proposition 6.1.5 and the discussion in Section 9.1 that $\mathcal{B}(\mathcal{H})$ is a Banach algebra. As we will see below, $\mathcal{B}(\mathcal{H})$ also carries an involution that makes it a **C*-algebra** (to be discussed in Chapter 11).

## 10.1. Adjoints

When we consider bounded operators over a Hilbert space, we are dealing in particular with bounded linear operators on a Banach space, so everything said in Chapter 9 applies. But we have further structure in the form of the inner product inducing the norm. This will imply several particular characteristics that we will address. It also implies the existence of the **adjoint** of an operator, in a sense that is almost but not exactly the same as the one we considered for operators on a Banach space. This existence will be deduced by an easy application of the Riesz Representation Theorem. We will explain more about this difference after we define the adjoint below.

***10.1.1. Lemma.***

*Let $T \in \mathcal{B}(\mathcal{H})$. Then the following statements are equivalent:*

*(i)* $T = 0$;

*(ii) for all* $\xi \in \mathcal{H}$, $\langle T\xi, \xi\rangle = 0$;

*(iii) for all* $\xi, \eta \in \mathcal{H}$, $\langle T\xi, \eta\rangle = 0$;

*Proof.* It is clear that (i) implies both (ii) and (iii). If we assume (iii), then for any $\xi \in \mathcal{H}$ we take $\eta = T\xi$, and we conclude that $\|T\xi\|^2 = 0$, so $T\xi = 0$. This shows (i). So it only remains to prove (ii) $\implies$ (iii): and this follows from the Polarization Identity. □

**10.1.2. Definition.**

*Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. An operator $S \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ is said to be an* ***adjoint*** *for $T$ if*

$$\langle T\xi, \eta\rangle = \langle \xi, S\eta\rangle, \quad \xi \in \mathcal{H}, \ \eta \in \mathcal{K}.$$

Most of the time we will consider adjoints in $\mathcal{B}(\mathcal{H})$. A straightforward application of Lemma 10.1.1 shows that the adjoint has to be unique (Exercise 10.1.1). Because of this, we are justified in using the standard notation $T^*$ for the adjoint of $T$. Since this is not exactly the Banach adjoint we defined before, we will use the notation $T^B$ in the few cases where we need to refer to the adjoint in the sense of Section 9.4.

*10.1.3. Remark.* The existence of the adjoint $T^*$ is not automatic for a bounded linear map on an inner product space. Indeed, consider $c_{00} \subset \ell^2(\mathbb{N})$. Let $T : c_{00} \to c_{00}$ be given by $(Ta)(k) = \sum_{j=k}^\infty \frac{a_j}{j^2}$. Then $T \in \mathcal{B}(c_{00})$, for $\|Ta\| \leq \|a\|$. Suppose we have $T^*$ as in Definition 10.1.2. Then

$$(T^*e_1)(j) = \langle T^*e_1, e_j\rangle = \langle e_1, Te_j\rangle = \overline{(Te_j)(1)} = \frac{1}{j^2}$$

for all $j$, so $T^*e_1$ cannot be in $c_{00}$.

We will soon prove that $T^*$ exists whenever $T \in \mathcal{B}(\mathcal{H})$—otherwise this section wouldn't exist, would it? But let us first talk about the difference between $T^*$ and $T^B$. They cannot possibly be equal, since for $\lambda \in \mathbb{C}$ we have $(\lambda T)^* = \bar{\lambda} T^*$, while $(\lambda T)^B = \lambda T^B$; that is, the map $T \longmapsto T^*$ is not linear, but rather it conjugates scalars. The Riesz Representation Theorem gives us an isomorphism $\Gamma : \mathcal{H} \to \mathcal{H}^*$ via $\Gamma(\xi) = \langle \cdot, \xi\rangle$. The issue is that $\Gamma$ is not linear, but **conjugate**

**linear**, i.e. $\Gamma(\eta+\lambda\xi)=\Gamma(\eta)+\bar\lambda\,\Gamma(\xi)$. So we have, for $\xi,\eta\in\mathcal{H}$,

$$(T^B\Gamma\xi)\eta=(\Gamma\xi)(T\eta)=\langle T\eta,\xi\rangle.$$

And

$$(\Gamma T^*\xi)\eta=\langle\eta,T^*\xi\rangle=\langle T\eta,\xi\rangle.$$

It follows that

$$T^B=\Gamma\,T^*\,\Gamma^{-1}. \tag{10.1}$$

In the general case, there is not much more to say. Because $\Gamma$ is a surjective isometry, $T^B$ and $T^*$ will have the same topological properties (closed range, bounded below, etc.) so we can apply all results from Chapter 9 to $T^*$.

In the finite-dimensional case, the relation between $T^B$ and $T^*$ can be made a bit more explicit. Consider $T:\mathbb{C}^n\to\mathbb{C}^n$ linear, and identify $(\mathbb{C}^n)^*$ with $\mathbb{C}^n$ in the canonical way. Then the Banach space adjoint of $T$ is the transpose $T^\top$, while the adjoint as defined above is the conjugate transpose $T^*$ (attention! This depends on us using an orthonormal basis; see Exercise 10.1.2). What the explicit nature of $\mathbb{C}^n$ gives us, is that it is trivial to identify its dual with itself, with an honest isomorphism (instead of conjugate-linear). Namely, we can see $y\in\mathbb{C}^n$ as a functional on $\mathbb{C}^n$ by

$$y'(x)=\sum_{k=1}^n x_ky_k.$$

This is exactly what we get if we think of the canonical basis as the dual basis of itself. Over the canonical basis,

$$(T^*)_{kj}=\langle T^*e_j,e_k\rangle=\langle e_j,Te_k\rangle=\overline{\langle Te_k,e_j\rangle}=\overline{T_{jk}}.$$

Meanwhile,

$$\begin{aligned}(T^By')(x)=y'(Tx)&=\sum_j y_j(Tx)_j\sum_j\sum_k y_jT_{jk}x_k\\&=\sum_k\sum_j T_{jk}y_j\,x_k=\sum_k(T^ty)_kx_k=(T^\top y)'(x).\end{aligned}$$

So $T^B=T^\top$.

### 10.1.4. Definition.

*A sesquilinear form $B:\mathcal{H}\times\mathcal{K}\to\mathbb{C}$ is **bounded** if there exists $c>0$ such that $|B(\xi,\eta)|\le c\,\|\xi\|\,\|\eta\|$ for all $\xi\in\mathcal{H}$ and $\eta\in\mathcal{K}$. The infimum of such numbers $c$ is denoted by $\|B\|$.*

**10.1.5. Proposition.**

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $\mathcal{H}_0 \subset \mathcal{H}$ and $\mathcal{K}_0 \subset \mathcal{K}$ dense subspaces. Let $B : \mathcal{H}_0 \times \mathcal{K}_0 \to \mathbb{C}$ be a bounded sesquilinear form. Then there exists a unique $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ such that $B(\xi, \eta) = \langle T\xi, \eta\rangle$ for all $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$, and $\|T\| = \|B\|$. If $\mathcal{K} = \mathcal{H}$ and $B$ is positive, so is $T$.*

*Proof.* Fix $\xi \in \mathcal{H}_0$ and define $\varphi : \mathcal{K}_0 \to \mathbb{C}$ by $\varphi(\eta) = \overline{B(\xi, \eta)}$. This is a linear functional, and

$$|\varphi(\eta)| = |B(\xi, \eta)| \leq \|B\| \, \|\xi\| \, \|\eta\|,$$

so bounded with $\|\varphi\| \leq \|B\| \, \|\xi\|$. It has a unique extension to $\mathcal{K}$ by Exercise 5.5.11 . By the Riesz Representation Theorem there exists $T\xi \in \mathcal{K}$ with $\overline{B(\xi, \eta)} = \langle \eta, T\xi\rangle$ for all $\eta \in \mathcal{K}$. That is,

$$B(\xi, \eta) = \langle T\xi, \eta\rangle, \qquad \eta \in \mathcal{K}.$$

The uniqueness in the Riesz Representation Theorem, together with the linearity of $B$ in the first coordinate, guarantee that the map $T : \xi \longmapsto T\xi$ is linear. From

$$|\langle T\xi, \eta\rangle| = |B(\xi, \eta))| \leq \|B\| \, \|\eta\| \, \|\xi\|$$

we get from Proposition 4.2.4 that $\|T\xi\| \leq \|B\| \, \|\xi\|$, showing that $T$ is bounded on $\mathcal{H}_0$, and $\|T\| \leq \|B\|$. By Proposition 6.1.9, applied to $T$ as map from $\mathcal{H}_0 \to \mathcal{H}$, $T$ extends to a bounded operator $T \in B(\mathcal{H}, \mathcal{K})$. The continuity of $T$ and $B$ guarantee that the relation $B(\xi, \eta) = \langle T\xi, \eta\rangle$ holds for all $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$.

Given $\varepsilon > 0$, choose unit vectors $\xi \in \mathcal{H}$, $\eta \in \mathcal{K}$ such that $|B(\xi, \eta)| + \varepsilon > \|B\|$. Then $\|B\| \leq \|T\| + \varepsilon$, and as $\varepsilon$ was arbitrary, we get the equality $\|T\| = \|B\|$.

The uniqueness of $T$ follows directly from the fact that if $\langle T\xi, \eta\rangle = 0$ for all $\xi, \eta$ the $T = 0$.

Finally, if $\mathcal{K} = \mathcal{H}$ and $B$ is positive, then $\langle T\xi, \xi\rangle = B(\xi, \xi) \geq 0$ for all $\xi \in \mathcal{H}$, so $T$ is positive. □

We are now in position to show that the adjoint always exists. It is not much harder to show the existence when $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$, but we will leave said proof to Exercise 10.4.6.

**10.1.6. Theorem.**

*Let $T \in \mathcal{B}(\mathcal{H})$. Then $T^*$ exists, and*

$$\|T^*\| = \|T\| = \|T^*T\|^{1/2}. \tag{10.2}$$

*Proof.* Applying Proposition 10.1.5 to the bounded sesquilinear form $[\xi,\eta] = \langle \xi, T\eta\rangle$, we obtain $S \in \mathcal{B}(\mathcal{H})$ with $\langle S\xi,\eta\rangle = \langle \xi, T\eta\rangle$ for all $\xi,\eta$; so $S = T^*$. Using Proposition 4.2.4,

$$\|T^*\xi\| = \sup\{|\langle T^*\xi,\eta\rangle| : \ \|\eta\| = 1\} = \sup\{|\langle \xi,T\eta\rangle| : \ \|\eta\| = 1\} \leq \|T\|\,\|\xi\|$$

for any $\xi \in \mathcal{H}$. Then $\|T^*\| \leq \|T\|$. Since it is clear from Definition 10.1.2 that $(T^*)^* = T$, we also have $\|T\| = \|(T^*)^*\| \leq \|T^*\|$, showing that $\|T^*\| = \|T\|$. Finally,

$$\begin{aligned}\|T\|^2 &= \sup\{\langle T\xi, T\xi\rangle : \ \xi \in \mathcal{H}, \ \|\xi\| = 1\} \\ &= \sup\{\langle T^*T\xi, \xi\rangle : \ \xi \in \mathcal{H}, \ \|\xi\| = 1\} \\ &\leq \|T^*T\|,\end{aligned}$$

and $\|T^*T\| \leq \|T^*\|\,\|T\| = \|T\|^2$. So $\|T\| = \|T^*T\|^{1/2}$ $\square$

The second equality in (10.2) is known as the **$C^*$-identity**. It plays a crucial role in the characterization of C$^*$-algebras in Chapter 11, and it will be used liberally and without reference.

As a converse to Theorem 10.1.6, the existence of the adjoint guarantees that $T$ is bounded. It is key that $T$ is defined everywhere, as unbounded densely defined selfadjoint operators are a thing (see Chapter F for an overview).

***10.1.7. Proposition.***

*Let $T : \mathcal{H} \to \mathcal{H}$ be linear. If the linear functional $f_\eta : \xi \longmapsto \langle T\xi, \eta\rangle$ is bounded for each $\eta \in \mathcal{H}$, then $T$ is bounded. Equivalently, if there exists $S : \mathcal{H} \to \mathcal{H}$ such that $\langle T\xi,\eta\rangle = \langle \xi, S\eta\rangle$ for all $\xi,\eta$, then $T$ is bounded.*

*Proof.* We use the Closed Graph Theorem. Suppose that $\{\xi_n\} \subset \mathcal{H}$ is a convergent sequence with $\{T\xi_n\}$ convergent. Let $\xi = \lim \xi_n$, and $\nu = \lim T\xi_n$. Then, for any $\eta$,

$$\langle \nu, \eta\rangle = \lim_n \langle T\xi_n, \eta\rangle = \lim_n f_\eta(\xi_n) = f_\eta(\xi) = \langle T\xi, \eta\rangle.$$

It follows that $\nu = T\xi$, so the graph of $T$ is closed and thus $T$ is bounded.

When $S : \mathcal{H} \to \mathcal{H}$ exists (as a function, it doesn't have to be neither linear nor bounded) such that $\langle T\xi,\eta\rangle = \langle \xi, S\eta\rangle$ for all $\xi,\eta \in \mathcal{H}$, then

$$|f_\eta(\xi)| = |\langle T\xi,\eta\rangle| = \langle \xi, S\eta\rangle| \leq \|S\eta\|\,\|\xi\|$$

and the previous part applies. $\square$

We summarize the basic properties of the operation of taking adjoint.

**10.1.8. Proposition.**

*Let $T, S \in \mathcal{B}(\mathcal{H})$, $\alpha \in \mathbb{C}$. Then*

(i) $(T + \alpha S)^* = T^* + \bar{\alpha}\, S^*$;

(ii) $(T^*)^* = T$;

(iii) $(TS)^* = S^* T^*$;

(iv) $T^* \in GL(\mathcal{H})$ *if and only if* $T \in GL(\mathcal{H})$; *in that case,* $(T^*)^{-1} = (T^{-1})^*$;

(v) $\sigma(T^*) = \{\bar{\lambda} : \ \lambda \in \sigma(T)\}$.

*Proof.* Exercise 10.3.3. □

*10.1.9. Remark.* Looking at Proposition 9.2.9 and the last statement in Proposition 10.1.8, one would be tempted to think that the mapping of the spectrum should also work with polynomials in $\mathbb{C}[z, \bar{z}]$. But this is not the case. Indeed, fix an orthonormal basis $\{\xi_j\}_{j\in\mathbb{N}}$ of $\mathcal{H}$ and consider the **shift operator** $S \in \mathcal{B}(\mathcal{H})$, induced by $S\xi_j = \xi_{j+1}$. Then $\sigma(S) = \overline{\mathbb{D}}$ (Example 9.5.10 in the case $p = 2$) and $S^*S = I$. So we have $\sigma(S^*S) = \{1\}$, but

$$\{|\lambda|^2 : \ \lambda \in \sigma(S)\} = [0, 1].$$

Given $T \in \mathcal{B}(\mathcal{H})$, two important subspaces, defined on Definition 6.1.1, will often be considered: the **kernel** of $T$ is the—closed, since $T$ is bounded—subspace $\ker T = \{\xi \in \mathcal{H} : \ T\xi = 0\}$, and **range** of $T$ is the—not necessarily closed—subspace $\operatorname{ran} T = \{T\xi : \ \xi \in \mathcal{H}\}$. We use $[\operatorname{ran} T]$ to denote the projection onto $\overline{\operatorname{ran}} T$.

When dealing with operators and their adjoints, it is very common to use the following relations between kernels and images of $T$ and $T^*$. The equalities (ii) and (iii) below should look familiar in light of Proposition 9.4.6, particularly when one notices that for Hilbert spaces the polar of a set is the orthogonal under the canonical identification of $\mathcal{H}$ with $\mathcal{H}^*$.

**10.1.10. Proposition.**

*Let $T \in \mathcal{B}(\mathcal{H})$. Then*

(i) $\ker T = \ker T^*T = \ker |T|$;

(ii) $\ker T = (\operatorname{ran} T^*)^\perp$;

*(iii)* $\overline{\operatorname{ran}} T = (\ker T^*)^\perp$;

*(iv)* $\overline{\operatorname{ran}} T^* = \overline{\operatorname{ran}} T^*T = \overline{\operatorname{ran}} |T|$.

*Proof.*

(i) The inclusion $\ker T \subset \ker T^*T$ is immediate. Now assume that $T^*T\xi = 0$. Then $\|T\xi\|^2 = \langle T\xi, T\xi\rangle = \langle T^*T\xi, \xi\rangle = 0$, so $\ker T^*T \subset \ker T$. We have $T^*T = |T|^2 = |T|^*|T|$, so the previous reasoning also shows that $\ker |T| = \ker T^*T$.

(ii) We have

$$\xi \in \ker T \iff \langle T\xi, \eta\rangle = 0 \ \forall \eta \iff \langle \xi, T^*\eta\rangle = 0 \ \forall\eta \iff \xi \in (\operatorname{ran} T^*)^\perp.$$

One could justify the converse of the first implication using Riesz's Representation Theorem (4.5.4) and that the dual separates points (Corollary 5.7.7), but it can proven simply by taking $\eta = T\xi$.

(iii) Applying the above to $T^*$ and taking orthogonal,

$$(\ker T^*)^\perp = (\operatorname{ran} T)^{\perp\perp} = \overline{\operatorname{ran} T}.$$

(iv) And then

$$\overline{\operatorname{ran} T^*} = (\ker T)^\perp = (\ker T^*T)^\perp = \overline{\operatorname{ran} (T^*T)^*} = \overline{\operatorname{ran} T^*T}.$$

Finally,

$$\overline{\operatorname{ran}} |T| = (\ker |T|)^\perp = (\ker T)^\perp = \overline{\operatorname{ran}} T^*.$$ □

Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. We know from Proposition 9.4.6 that $\operatorname{ran} T$ is closed if and only if $\operatorname{ran} T^*$ is closed. Here is an argument specific for operators on Hilbert spaces. First of all, because $T^{**} = T$, only one implication needs to be proven.

Suppose $\operatorname{ran} T$ is closed. Let $H_1 = (\ker T)^\perp$ and $H_2 = \operatorname{ran} T$. As an operator $T : H_1 \to H_2$, $T$ is invertible by the Inverse Mapping Theorem 6.3.6. Let $i_1 : H_1 \to \mathcal{H}$ and $i_2 : H_2 \to \mathcal{H}$ be the inclusions. Then $i_2^* T i_1 : H_1 \to H_2$ is invertible and thus its adjoint is also invertible. The adjoint is $i_1^* T^* i_2 : H_2 \to H_1$. The range of $T^*$ is a dense subspace of $H_1$, since $\overline{\operatorname{ran} T^*} = (\ker T)^\perp$; so $i_1^* T^* = T^*$ by Exercise 10.1.6. As $T^* i_2 = i_1^* T^* i_2$ is onto $H_1$, we conclude $\operatorname{ran} T^* = H_1$, closed.

**10.1.11. Definition.**

*An operator $U \in \mathcal{B}(\mathcal{H})$ is a* ***unitary*** *if $U^*U = UU^* = I$. The multiplicative group of unitaries in $\mathcal{B}(\mathcal{H})$ is denoted by $\mathcal{U}(\mathcal{H})$.*

*An operator $V \in \mathcal{B}(\mathcal{H})$ such that $V^*V = I$ is an* ***isometry****. When $VV^* = I$ we say that $V$ is a* ***co-isometry****. An isometry that is not a unitary is called a* ***proper isometry****.*

We saw unitaries in Chapter 4, in that they are the homomorphisms (and necessarily, isomorphisms) of Hilbert spaces. In finite dimension each of the equalities $U^*U = I$ and $UU^* = I$ implies each other (Exercise 10.1.4). In infinite dimension the unilateral shift is the canonical example of an operator satisfying $S^*S = I$ but $SS^* \neq I$; that is, $S$ is an proper isometry. It is trivial from the definition that $U$ is a unitary if and only if $U^*$ is a unitary. Also, using the C*-identity,

$$\|U\| = \|U^*U\|^{1/2} = \|I\|^{1/2} = 1.$$

For any isometry $V$ (in particular, for a unitary), we have

$$\|V\eta\|^2 = \langle V\eta, V\eta\rangle = \langle V^*V\eta, \eta\rangle = \langle \eta, \eta\rangle = \|\eta\|^2, \qquad \eta \in \mathcal{H},$$

hence the name. Then, for any $T \in \mathcal{B}(\mathcal{H})$,

$$\|VT\| = \sup\{\|VT\xi\| : \ \|\xi\| = 1\} = \sup\{\|T\xi\| : \ \|\xi\| = 1\} = \|T\|.$$

Taking adjoints, or with a similar argument as above, we also have $\|TV^*\| = \|T\|$ for all $T \in \mathcal{B}(\mathcal{H})$.

***10.1.12. Proposition.***

*Let $U \in \mathcal{B}(\mathcal{H})$ be a unitary. Then $\sigma(U) \subset \mathbb{T}$.*

*Proof.* Since $\|U\| = 1$, we know that $\sigma(U) \subset \overline{\mathbb{D}}$. Suppose $|\lambda| < 1$. Then

$$\|I - (I - \lambda U^*)\| = \|\lambda U^*\| = |\lambda| < 1.$$

By Lemma 6.2.1, $I - \lambda U^*$ is invertible; and then $U - \lambda I = U(I - \lambda U^*)$ is invertible. Thus $\lambda \notin \sigma(U)$. It follows that $\sigma(U) \subset \mathbb{T}$. □

With more knowledge about $\mathcal{B}(\mathcal{H})$ it will be easy to see that any compact $K \subset \mathbb{T}$ can be the spectrum of a unitary (see page 693).

### 10.1.1. Exercises.

(10.1.1) Prove the uniqueness of the adjoint of $T \in \mathcal{B}(\mathcal{H})$.

(10.1.2) Let $\mathcal{H} = \mathbb{C}^2$. Let $\{\xi_1, \xi_2\}$ be the canonical basis, and $T$ the operator induced $T\xi_1 = 0$, $T\xi_2 = \xi_1$.

(i) Find the matrix form of $T$ and $T^*$ with respect to the canonical basis. Confirm that $T^*$ is the conjugate transpose.

(ii) Find the matrix form of $T$ and $T^*$ with respect to the basis $\eta_1 = \xi_1$, $\eta_2 = \xi_1 + \xi_2$. Is the matrix of $T^*$ the conjugate transpose of the matrix of $T$? Why?

(10.1.3) Let $\mathcal{H}$ a Hilbert space and $\mathcal{H}_0 \subset \mathcal{H}$ a dense subspace. Let $B : \mathcal{H}_0 \times \mathcal{H}_0 \to \mathbb{C}$ be a bounded sesquilinear form. Show that $B$ admits a unique extension to a bounded sesquilinear form $\tilde{B} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, with $\|\tilde{B}\| = \|B\|$.

(10.1.4) Let $T \in \mathcal{B}(\mathcal{H})$, where $\dim \mathcal{H} < \infty$. Show that $T^*T = I$ if and only if $TT^* = I$.

(10.1.5) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces of the same dimension, and $\{\xi_j\}_{j\in J}$ and $\{\eta_j\}_{j\in J}$ orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$ respectively. Show that the assignment $U : \xi_j \longmapsto f_j$ induces a unique bounded linear operator $U \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and that $U$ is a unitary.

(10.1.6) Let $K \subset \mathcal{H}$ be a closed subspace and $i : K \to \mathcal{H}$ the inclusion. Show that $i^*$ is the orthogonal projection onto $K$.

(10.1.7) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta \in \mathcal{H}$ with $\|\xi\| = \|\eta\|$. Show that there exists a unitary $U \in \mathcal{B}(\mathcal{H})$ such that $U\xi = \eta$.

(10.1.8) Let $\mathcal{H}$ be a Hilbert space, and consider the direct sum $\mathcal{H} \oplus \mathcal{H}$. But instead of considering the natural Hilbert space norm $\|(\xi, \eta)\| = (\|\xi\|^2 + \|\eta\|^2)^{1/2}$, we consider the norm $\|(\xi, \eta)\|_1 = \|\xi\| + \|\eta\|$. Since these two norms are equivalent, $\mathcal{H} \oplus_1 \mathcal{H}$ (with $\|\cdot\|_1$) is a Banach space. Show that there exist elements $(\xi, \eta), (\xi', \eta') \in \mathcal{H} \oplus_1 \mathcal{H}$ with $\|(\xi, \eta)\| = \|(\xi', \eta')\|$ but such that no linear isometry maps $(\xi, \eta) \longmapsto (\xi', \eta')$ (compare with Exercise 7.5.19).

## 10.2. Numerical Range and Numerical Radius

We will briefly introduce the numerical range and prove the most important result about it, the Toeplitz–Hausdorff Theorem. We mention, tough, that we are not even scratching the surface here. The study of the numerical range is itself an entire area of mathematics, with even regular conferences on the topic.

**10.2.1. Definition.**

*Given $T \in \mathcal{B}(\mathcal{H})$, the* ***numerical range*** *of $T$ is the set*

$$W(T) = \{\langle T\xi, \xi\rangle : \ \xi \in \mathcal{H}, \ \|\xi\| = 1\}.$$

*The* ***numerical radius*** *of $T$ is the number*

$$\omega(T) = \sup\{|\lambda| : \ \lambda \in W(T)\}.$$

The numerical radius is actually a norm, equivalent to the operator norm (Exercise 10.2.2). Since $T$ is bounded, it is clear that $\overline{W(T)}$ is a compact subset of $\mathbb{C}$. Less obvious is that $W(T)$ is convex. This is a celebrated theorem, and over the years since the original proof in 1918, simpler and simpler arguments have been found.

***10.2.2. Theorem.*** ***(Toeplitz–Hausdorff 1918)***

*Let $T \in B(H)$. Then $W(T)$ is convex.*

*Proof.* Let $\lambda, \mu \in W(T)$ be distinct. Take $\xi_1, \xi_2 \in \mathcal{H}$ with $\|\xi_1\| = \|\xi_2\| = 1$ and $\lambda = \langle T\xi_1, \xi_1\rangle \neq \langle T\xi_2, \xi_2\rangle = \mu$. We know that $\xi_1$ and $\xi_2$ are not colinear: if $\xi_2 = \alpha\xi_1$, then $|\alpha| = 1$ and we would have $\lambda = \mu$. So, $\xi_1 + t\xi_2 \neq 0$ for all scalars $t$. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(t) = \frac{\langle T(\xi_1 + t\xi_2), \xi_1 + t\xi_2\rangle - \lambda \langle \xi_1 + t\xi_2, \xi_1 + t\xi_2\rangle}{(\mu - \lambda) \langle \xi_1 + t\xi_2, \xi_1 + t\xi_2\rangle}.$$

This function is continuous, since it is a rational function on $t$ and the denominator is never zero. We clearly have $f(0) = 0$, and $\lim_{t\to\infty} f(t) = 1$. Then, given an arbitrary $\alpha \in (0, 1)$, there exists $t_\alpha$ with $f(t_\alpha) = \alpha$. If we write $\xi_\alpha = \xi_1 + t_\alpha\xi_2$, we have that

$$\alpha\mu+(1-\alpha)\lambda = \alpha(\mu-\lambda)+\lambda = f(t_\alpha)(\mu-\lambda)+\lambda = \frac{1}{\|\xi_\alpha\|^2} \langle T\xi_\alpha, \xi_\alpha\rangle \in W(T). \quad \square$$

Since $\sigma(T)$ is compact, we know from Exercise 7.5.27 that $\operatorname{conv}\sigma(T)$ is compact.

***10.2.3. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$. Then* $\operatorname{conv}\sigma(T) \subset \overline{W(T)}$.

*Proof.* Let $\lambda \in \sigma(T)$. If $\lambda \in \sigma_{ap}(T)$, there exists a sequence $\{\xi_n\} \subset \mathcal{H}$ with $\|\xi_n\| = 1$ for all $n$ and $T\xi_n - \lambda\xi_n \to 0$. Then

$$\lambda = \lim_n \langle \lambda\xi_n, \xi_n\rangle = \lim_n \langle T\xi_n, \xi_n\rangle \in \overline{W(T)}.$$

If $\lambda \in \sigma_r(T)$, then $\overline{\operatorname{ran}(T - \lambda I)} \subsetneq \mathcal{H}$. So $\ker(T^* - \bar{\lambda} I) = \operatorname{ran}(T - \lambda I)^\perp \neq \{0\}$; that is, there exists $\xi \in \mathcal{H}$ with $T^*\xi = \bar{\lambda}\xi$. Then

$$\lambda = \langle \xi, \bar{\lambda}\xi\rangle = \langle \xi, T^*\xi\rangle = \langle T\xi, \xi\rangle \in W(T).$$

So $\sigma(T) \subset \overline{W(T)}$. As $W(T)$ is convex (Theorem 10.2.2) we get $\operatorname{conv}\sigma(T) \subset \overline{W(T)}$. ☐

The inclusion in Proposition 10.2.3 becomes an equality when $T$ is normal; we will prove this in Corollary 12.4.2, and a bit earlier (Exercise 10.6.24) for normal compact operators. Such proof requires more machinery (concretely the Spectral Theorem) and that is why it has to wait.

The closure in Proposition 10.2.3 cannot be dispensed in general. This will follow from the following (where, ironically, no closures are needed):

***10.2.4. Lemma.***

*Let $\{\xi_n\} \subset \mathcal{H}$ be an orthonormal basis, $t \in \ell^\infty(\mathbb{N})$, and $T \in \mathcal{B}(\mathcal{H})$ given by $T\xi_n = t_n\xi_n$ and extended by linearity and continuity (such $T$ is a **diagonal operator**). Then*

$$W(T) = \operatorname{conv}\{t_n : \ n\}.$$

*Proof.* Consider a convex combination $\sum_{j=1}^m \alpha_j\, t_{n_j}$. Let $\xi = \sum_{j=1}^m \alpha_j^{1/2}\, \xi_{n_j}$. Then $\|\xi\| = 1$ and

$$\langle T\xi, \xi\rangle = \sum_{j=1}^m \alpha_j\, t_{n_j}$$

So $\operatorname{conv}\{t_n\} \subset W(T)$. For $\lambda \in W(T) \setminus \operatorname{conv}\{t_n\}$, by the convexity and boundedness of $\operatorname{conv}\{t_n\}$ there exists a line through $\lambda$, with $\operatorname{conv}\{t_n\}$ on one side of the line (possibly touching it). As $\alpha W(T) + \beta = W(\alpha T + \beta I)$ (Exercise 10.3.1) and $\alpha\sigma(T)+\beta = \sigma(\alpha T+\beta I)$ (Exercise 9.5.1) for all $\alpha, \beta \in \mathbb{C}$, we may assume without loss of generality that $\lambda = 0$ and that $\operatorname{conv}\{t_n\}$ is contained in the closed upper half plane. The fact that $0 \in W(T)$ gives us $\xi = \sum_j c_j\xi_j$ with $\|\xi\| = 1$ and $0 = \langle T\xi, \xi\rangle$; then $\sum_n |c_n|^2 = 1$ and $0 = \sum_n t_n|c_n|^2$. As $\operatorname{Im} t_n \geq 0$ for all $n$ and $|c_n|^2 \geq 0$, whenever $\operatorname{Im} t_n > 0$ we have $c_n = 0$. By removing those indices, we may assume that $t_n \in \mathbb{R}$ for all $n$. If not all $t_n$ are zero, to make the sum zero there has to exist $n_1, n_2$ with $t_{n_1} < 0$ and $t_{n_2} > 0$; in either case, $0 \in \operatorname{conv}\{t_n\}$. So $W(T) = \operatorname{conv}\{t_n\}$. ☐

*10.2.5. Example* . Let $\{t_n\}$ be a dense sequence in $(0, 1)$ and $T$ a diagonal operator with diagonal $\{t_n\}$. Then $\sigma(T) = [0, 1]$ (see Example 9.5.6). By Lemma 10.2.4, $W(T) = \operatorname{conv}\{t_n\} = (0, 1)$. So $\sigma(T) \subsetneq W(T)$, showing that the closure cannot be avoided in Proposition 10.2.3. By including 0 and/or 1 in the set $\{t_n\}$ we can get $W(T)$ to be any of $(0, 1)$, $(0, 1]$, $[0, 1)$, and $[0, 1]$.

**10.2.1. Exercises.**

(10.2.1) Let $T \in \mathcal{B}(\mathcal{H})$. Show that $W(T^*) = \{\bar{\lambda} : \lambda \in W(T)\}$, and $\omega(T) = \omega(T^*)$.

(10.2.2) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Show that $\omega(T)$ defines a norm on $\mathcal{B}(\mathcal{H})$, equivalent to the operator norm; concretely,

$$\omega(T) \leq \|T\| \leq 2\omega(T).$$

The second inequality requires Proposition 10.3.3 and a basic knowledge of selfadjoint operators.

(10.2.3) Let $B \in \mathcal{B}(\mathcal{H})^{\mathrm{sa}}$ such that $\|I_{\mathcal{H}} + iB\| \leq 1$. Show that $B = 0$.

## 10.3. Selfadjoint Operators

Selfadjoint operators behave in $\mathcal{B}(\mathcal{H})$, from certain points of view, as the real numbers do within $\mathbb{C}$. They are easier (which doesn't mean "easy"!) to deal with, there specific results about them, and sometimes one is able to reduce an argument about general operator to an argument about selfadjoints.

**10.3.1. Definition.**

*We say that $T \in \mathcal{B}(\mathcal{H})$ is* ***normal*** *if $T^*T = TT^*$ and it is* ***selfadjoint*** *(or* ***Hermitian****) if $T^* = T$.*

It is a straightforward observation that a selfadjoint operator is normal. Normal operators are the "good guys", as we will see. For Lemma 10.3.2, we remark that it is perfectly possible to have $\sigma_p(T) = \varnothing$ for $T$ normal; we saw examples of this in Example 9.5.7.

***10.3.2. Lemma.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Then* $\ker T^* = \ker T$. *In particular,* $\sigma_p(T^*) = \sigma_p(T)$.

*Proof.* We have

$$\|T^*\xi\|^2 = \langle TT^*\xi, \xi\rangle = \langle T^*T\xi, \xi\rangle = \|T\xi\|^2,$$

so $\ker T^* = \ker T$. As $T$ is normal if and only if $T - \lambda I$ is normal (Exercise 10.3.2) and $(T - \lambda I)^* = T^* - \bar{\lambda} I$,

$$\begin{aligned}\lambda \in \sigma_p(T) &\iff \ker(T - \lambda I) \neq \{0\} \iff \ker(T^* - \bar{\lambda} I) \neq \{0\}\\ &\iff \bar{\lambda} \in \sigma_p(T^*).\end{aligned}$$

□

Here are some nice properties of normal operators.

***10.3.3. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal . Then*

*(i)* $\sigma(T) = \sigma_{ap}(T)$;

*(ii)* $\operatorname{spr}(T) = \|T\|$;

*(iii)* $\|T\| = \sup\{|\langle T\xi, \xi\rangle| : \ \|\xi\| = 1\}$;

*If moreover $T$ is selfadjoint, then*

*(iv)* $\sigma(T) \subset \mathbb{R}$.

*Proof.*

(i) Since $T - \lambda I$ is normal, Lemma 10.3.2 and Proposition 10.1.10 imply that

$$\ker(T - \lambda I) = \ker(T - \lambda I)^* = (\operatorname{ran}(T - \lambda I))^{\perp}.$$

So if $\lambda \in \sigma(T)$ and $\ker(T - \lambda I) = \{0\}$, then $\operatorname{ran} T - \lambda I$ is dense and hence $\lambda \notin \sigma_r(T)$. As $\sigma(T) = \sigma_{ap}(T) \cup \sigma_r(T)$ (Exercise 9.5.11) and we have just shown that $\sigma_r(T) = \varnothing$, we have that $\sigma(T) = \sigma_{ap}(T)$.

(ii) Let $S = T^*T$. Then $S$ is selfadjoint. From $\|S^2\| = \|S^*S\| = \|S\|^2$, we can get by induction that $\|S^{2^n}\| = \|S\|^{2^n}$. Using Theorem 9.2.16 and that $T$ is normal,

$$\operatorname{spr}(T) = \lim_n \|T^n\|^{1/n} = \lim_n \|(T^n)^*T^n\|^{1/2n} = \lim_n \|(T^*T)^n\|^{1/2n}$$

$$= \lim_n \|(T^*T)^{2n}\|^{1/4n} = \|T^*T\|^{1/2} = \|T\|.$$

(iii) By Proposition 10.2.3 we have that

$$\|T\| = \operatorname{spr}(T) \leq \sup \overline{\operatorname{conv}} \sigma(T) \leq \sup\{|\lambda| : \lambda \in W(T)\} \leq \|T\|.$$

So $\|T\| = \sup\{|\langle T\xi, \xi\rangle| : \|\xi\| = 1\}$.

(iv) Now we assume that $T$ is selfadjoint. For any $\xi \in \mathcal{H}$,

$$\langle T\xi, \xi\rangle = \langle \xi, T^*\xi\rangle = \langle \xi, T\xi\rangle = \overline{\langle T\xi, \xi\rangle},$$

so $\langle T\xi, \xi\rangle \in \mathbb{R}$ and so $W(T) \subset \mathbb{R}$. By Proposition 10.2.3, we get that $\sigma(T) \subset \mathbb{R}$.

□

*10.3.4. Remark*. The proof of the equality $\|T\| = \sup\{|\langle T\xi, \xi\rangle| : \|\xi\| = 1\}$ for a normal operator uses the Toeplitz–Hausdorff Theorem, which is a non-trivial result (even if the proof we used is not that complicated, it is not obvious either). In the case where $T$ is selfadjoint, here is an elementary argument that appears in [Con90]. Let $c = \sup\{|\langle T\xi, \xi\rangle| : \|\xi\| = 1\}$. It is trivial that $c \leq \|T\|$, so we focus on the reverse inequality. Fix $\xi, \eta \in \mathcal{H}$ with $\|\xi\| = \|\eta\| = 1$. Then, using that $T^* = T$,

$$\langle T(\xi \pm \eta), \xi \pm \eta\rangle = \langle T\xi, \xi\rangle + \langle T\eta, \eta\rangle \pm 2\mathrm{Re}\,\langle T\xi, \eta\rangle,$$

where the signs are coordinated. Subtracting the two equations above (the one with minus from the one with plus) we get

$$\langle T(\xi + \eta), \xi + \eta\rangle - \langle T(\xi - \eta), \xi - \eta\rangle = 4\mathrm{Re}\,\langle T\xi, \eta\rangle.$$

Since $|\langle T\nu, \nu\rangle| \leq c\,\|\nu\|^2$ for any $\nu \in \mathcal{H}$, using the Parallelogram Identity

$$4\mathrm{Re}\,\langle T\xi, \eta\rangle \leq c(\|\xi + \eta\|^2 + \|\xi - \eta\|^2) = 2c(\|\xi\|^2 + \|\eta\|^2) = 4c.$$

If we write $|\langle T\xi, \eta\rangle| = \lambda\,\langle T\xi, \eta\rangle$ for an appropriate $\lambda \in \mathbb{T}$, we obtain

$$\langle T\xi, \eta\rangle \leq c, \qquad \xi, \eta \in \mathcal{H},\ \|\xi\| = \|\eta\| = 1.$$

Putting $\eta = \frac{1}{\|T\xi\|} T\xi$ this becomes $\|T\xi\| \leq c$ for all $\xi$ with $\|\xi\| = 1$. Then $\|T\| \leq c$.

*10.3.5. Remark*. We showed in Proposition 10.3.3 that that $\sigma(T) \subset \mathbb{R}$ if $T = T^*$ by appealing to previous knowledge about the numerical rank. Here is an argument that does not use the numerical range.

Let $\lambda \in \sigma(T)$, and consider the operator $e^{iT}$; we can define this via the series for the exponential, and the operator we obtain agrees with the one we would obtain via Riesz Functional Calculus as in Theorem 9.3.2 (Exercise 9.3.2). By the

Spectral Mapping Theorem (Theorem 9.3.3), $e^{i\lambda} \in \sigma(e^{iT})$. By working with the exponential series and the continuity of taking adjoints, $(e^{iT})^* = e^{-iT^*} = e^{-iT}$. Then $(e^{iT})^* e^{iT} = e^{-iT} e^{iT} = I$, and $e^{iT}(e^{iT})^* = e^{iT} e^{-iT} = I$, so $e^{iT}$ is a unitary. By Proposition 10.1.12, $e^{i\lambda} \in \mathbb{T}$, and so $\lambda \in \mathbb{R}$.

It is possible for non-selfadjoint operators to have real spectrum. For instance, if we take $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in M_2(\mathbb{C})$, then $\sigma(T) = \{1\}$ but $T \neq T^*$. Another example is the Volterra operator (subsection 10.6.7). When a normal operator has real spectrum, it has to be selfadjoint (Proposition 10.3.7).

As opposed to the spectrum, the numerical range does contain enough information to tell if $T$ is selfadjoint:

***10.3.6. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$. The following statements are equivalent:*

*(i) $\langle T\xi, \xi\rangle \in \mathbb{R}$ for all $\xi \in \mathcal{H}$;*

*(ii) $T \in \mathcal{B}(\mathcal{H})^{\mathrm{sa}}$.*

*Proof.* Assume first that $\langle T\xi, \xi\rangle \in \mathbb{R}$ for all $\xi \in \mathcal{H}$. We have
$$\langle T\xi, \xi\rangle = \overline{\langle T\xi, \xi\rangle} = \langle \xi, T\xi\rangle = \langle T^*\xi, \xi\rangle.$$
So $\langle (T - T^*)\xi, \xi\rangle = 0$ for all $\xi$. By the Polarization Identity (Proposition 4.2.6), applied to the sesquilinear form $\psi(\xi, \eta) = \langle (T - T^*)\xi, \eta\rangle$, we have $\langle (T - T^*)\xi, \eta\rangle = 0$ for all $\xi, \eta$, and so $T - T^* = 0$.

Conversely, if $T = T^*$ then
$$\langle T\xi, \xi\rangle = \langle \xi, T\xi\rangle = \overline{\langle T\xi, \xi\rangle},$$
so $\langle T\xi, \xi\rangle \in \mathbb{R}$. ☐

Let us see some examples of normal and selfadjoint operators in Chapter 12, but let us check a few now.

As constructed in Example 9.5.6 we can fix an orthonormal basis $\{\xi_n\}$ on $\mathcal{H}$, $g \in \ell^\infty(\mathbb{N})$, and define $T : \mathcal{H} \to \mathcal{H}$ induced by $T\xi_n = g(n)\,\xi_n$. This is a **diagonal operator**, that satisfies $\|T\| = \|g\|_\infty$, $\sigma(T) = \sigma_{ap}(T) = \overline{g(\mathbb{N})}$, and it is normal (Exercise 10.3.6).

Now we can do the following. Given any $K \subset \mathbb{C}$ compact, let $\{r_n\} \subset K$ be dense in $K$. For the diagonal operator $T$ given by the sequence $\{r_n\}$, we get that $\sigma(T) = K$. That is, any compact subset of $\mathbb{C}$ appears as the spectrum of some normal operator.

Diagonal operators as above will always have lots of eigenvalues, because every element in the prescribed orthonormal basis will be an eigenvector. There exist normal and selfadjoint operators with no eigenvalues, though. The canonical example is $T \in B(L^2[0,1])$, with $T$ the multiplication operator given by

$$(Tf)(t) = tf(t). \tag{10.3}$$

Then $\sigma(t) = [0,1]$, $T$ is injective, and $\sigma_p(T) = \varnothing$ (Exercise 10.3.8). What about other Hilbert spaces, like $\ell^2(\mathbb{Z})$? Diagonal operators are very natural to produce on $\ell^2(\mathbb{Z})$, but can we do something like (10.3)? Well, all separable infinite-dimensional Hilbert spaces are isomorphic! And the isomorphism $\Gamma : \ell^2(\mathbb{Z}) \to L^2[0,1]$ can be made explicit, as all we need to do is map one orthonormal basis onto another. So for instance we can take the canonical basis $\{e_n\}$ on $\ell^2(\mathbb{Z})$ and have

$$\Gamma(e_n) = g_n, \qquad \text{where} \qquad g_n(t) = e^{2\pi i n t}.$$

Calculating the Fourier coefficients for $T$, we get that

$$T_{nm} = \langle Tg_m, g_n\rangle = \int_0^1 t\, e^{2\pi i(m-n)t}\, dt = -\frac{1}{2\pi i(m-n)}$$

when $n \neq m$, and $T_{nn} = \frac{1}{2}$ for all $n$. So the operator $T' \in B(\ell^2(\mathbb{Z}))$ induced by

$$T'e_m = \frac{1}{2}\, e_m - \sum_{n\in\mathbb{Z},\ n\neq m} \frac{1}{2\pi i(m-n)}\, e_n$$

satisfies that it is selfadjoint, injective, $\sigma(T') = [0,1]$, $\sigma_p(T') = \varnothing$.

We learned the proof below from Ryszard Szwarc. See Remark 12.4.3 for a less creative straightforward proof at the cost of using machinery.

***10.3.7. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal with $\sigma(T) \subset \mathbb{R}$. Then $T = T^*$.*

*Proof.* Since $i \notin \sigma(T)$, we have that $iI \pm T$ is invertible. Define

$$V = (iI - T)(iI + T)^{-1}.$$

If $|\lambda| \neq 1$,

$$\begin{aligned} V - \lambda I &= \big[(iI - T) - \lambda(iI + T)\big]\,(iI + T)^{-1} \\ &= \big[i(1-\lambda) - (1+\lambda)T\big]\,(iI + T)^{-1} \\ &= (1+\lambda)\left[\frac{i(1-\lambda)}{1+\lambda} - T\right](iI + T)^{-1}. \end{aligned}$$

Writing $\lambda = re^{it}$,

$$\frac{i(1-\lambda)}{1+\lambda} = \frac{-2r\sin t + i(1-r^2)}{1 + r^2 + 2r\cos t} \notin \mathbb{R}$$

since $r \neq 1$. So $V - \lambda I$ is invertible. When $\lambda = -1$ we have $V - \lambda I = 2i(iI+T)^{-1}$, invertible. Thus

$$\sigma(V) \subset \mathbb{T} \setminus \{-1\}.$$

Since $V$ is normal, by Proposition 10.3.3 $\|V\| = 1$. As $V^{-1}$ is also normal with $\sigma(V^{-1}) \subset \mathbb{T}$ (by Proposition 9.2.9), $\|V^{-1}\| = 1$. Then

$$\|\xi\| = \|V^{-1}V\xi\| \leq \|V\xi\| \leq \|\xi\|.$$

So $V$ is isometric. Being invertible, it is a unitary. One checks algebraically that

$$T = i(I-V)(I+V)^{-1}.$$

Then, using that $V^* = V^{-1}$,

$$\begin{aligned} T^* &= -i(I+V^{-1})^{-1}(I-V^{-1}) = -i(V+I)^{-1}(V-I) \\ &= i(I-V)(I+V)^{-1} = T. \end{aligned}$$

□

*10.3.8. Remark*. Using a few more tools, the proof of Proposition 10.3.7 can be made simpler; this argument was mentioned to us by Ruy Exel. Note that if $T$ is normal, then $T + icI$ is normal for all $c \in \mathbb{R}$. Fix $\xi \in \mathcal{H}$ with $\|\xi\| = 1$. Write $\langle T\xi, \xi\rangle = a + ib$. Then, using Proposition 10.3.3 and that $\sigma(T) \subset \mathbb{R}$,

$$\begin{aligned} a^2 + b^2 + 2bc + c^2 = |a + i(b+c)|^2 &= |\langle (T+ic)\xi, \xi\rangle| \\ &= \sup\{|\lambda|^2 :\ \lambda \in \sigma(T+icI)\| = \sup\{|\lambda - ic|^2 :\ \lambda \in \sigma(T)\} \\ &= \sup\{\lambda^2 + c^2 :\ \lambda \in \sigma(T)\} = \|T\|^2 + c^2. \end{aligned}$$

It follows that $a^2 + b^2 + 2bc \leq \|T\|^2$ for all $c \in \mathbb{R}$; this forces $b = 0$, so $\langle T\xi, \xi\rangle \in \mathbb{R}$, and then Proposition 10.3.6 implies that $T$ is selfadjoint.

**10.3.9. Definition.**

*Let $T \in \mathcal{B}(\mathcal{H})$. The **real part** and the **imaginary part** of $T$ are the selfadjoint operators*

$$\operatorname{Re} T = \frac{T + T^*}{2}, \qquad \operatorname{Im} T = \frac{T - T^*}{2i}.$$

So every $T$ can be written as $T = H + iK$, with $H, K$ selfadjoint. In particular, the selfadjoint operators span $\mathcal{B}(\mathcal{H})$.

### 10.3.1. Exercises.

(10.3.1) Let $T \in \mathcal{B}(\mathcal{H})$ and $\alpha, \beta \in \mathbb{C}$. Show that $W(\alpha T + \beta I) = \alpha W(T) + \beta$.

(10.3.2) Let $T \in \mathcal{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$. Show that $T$ is normal if and only if $T - \lambda I$ is normal.

(10.3.3) Prove Proposition 10.1.8.

(10.3.4) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\xi \in \mathcal{H}$ and $\mathcal{H}_1 \subset \mathcal{H}$ be the subspace

$$\mathcal{H}_1 = \overline{\{p(T,T^*)\xi : \ p \in \mathbb{C}[x,y]\}}.$$

Show that $\mathcal{H}_1^\perp$ is invariant for both $T$ and $T^*$.

(10.3.5) Recall that an operator $T \in \mathcal{B}(\mathcal{H})$ is said to be **bounded below** if there exists $k > 0$ with $\|T\xi\| \geq k\|\xi\|$ for every $\xi \in \mathcal{H}$. Show that the following statements are equivalent:

(i) $T$ is bounded below;

(ii) $T$ admits a left inverse;

(iii) $T$ is injective and has closed range.
*(Hint: use the Inverse Mapping Theorem).*

(10.3.6) Let $T$ be a diagonal operator. Find $T^*$ and show that $T$ is normal.

(10.3.7) Show that a diagonal operator $T$ is selfadjoint if and only if all its diagonal entries are real.

(10.3.8) Let $T$ as in (10.3). Show that $T$ is injective, selfadjoint, with $\sigma(T) = [0,1]$ and $\sigma_p(T) = \varnothing$.

(10.3.9) Let $\mathcal{H}$ be a Hilbert space and $\xi, \eta \in \mathcal{H}$.

(i) Show that it is possible to choose $\xi, \eta$ in such a way that no selfadjoint $T$ satisfies $T\xi = \eta$, even if both are nonzero.

(ii) Show that there exists $T$ normal with $T\xi = \eta$.

## 10.4. Positive operators

We have mentioned the idea that to a certain extent one can think that the selfadjoint elements in $\mathcal{B}(\mathcal{H})$ play a similar role to that of $\mathbb{R}$ in $\mathbb{C}$, and that the adjoint can be thought of as the conjugate; with many differences, born of the fact that multiplication in $\mathcal{B}(\mathcal{H})$ is not commutative in general. In that case, one could think that the analog of the non-negative numbers would be played by the operators $T^*T$ (in analogy with $|z|^2 = z\bar{z}$). Such operators would satisfy $\langle T^*T\xi, \xi\rangle = \|T\xi\|^2 \geq 0$. We will use this latter property as our starting point.

### 10.4.1. Definition Basic Properties.

**10.4.1. Definition.**

*An operator $T \in \mathcal{B}(\mathcal{H})$ is* ***positive*** *if $\langle T\xi, \xi\rangle \geq 0$ for all $\xi \in \mathcal{H}$. We will use the notation $T \geq 0$.*

*The order between selfadjoint operators $A, B \in \mathcal{B}(\mathcal{H})^{\mathrm{sa}}$ is given by saying that $A \leq B$ when $B - A \geq 0$.*

*A* ***conjugate*** *of $A \in \mathcal{B}(\mathcal{H})$ is $TAT^*$ for some $T \in \mathcal{B}(\mathcal{H})$.*

The order between selfadjoint operators is a partial order, as not all pairs of selfadjoint operators are comparable. The condition $A \geq 0$ says that $A$ is positive.

It is worth mentioning that we showed in Proposition 6.3.15 that a positive operator is necessarily bounded. The reader will hopefully translate the argument to the context of Hilbert spaces in Exercise 10.4.1.

It follows from Proposition 10.3.6 that positive operators are selfadjoint. This uses in an essential way that the underlying field is $\mathbb{C}$. Indeed, if we consider real spaces, positive operators exist that are not selfadjoint. Easiest example is probably

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \mathcal{B}(\mathbb{R}^2);$$

for any $x \in \mathbb{R}^2$, we have $\langle Tx, x\rangle = x_1^2 + x_2^2 + x_1x_2 \geq (|x_1| - |x_2|)^2 \geq 0$. So $T$ is positive but not selfadjoint.

Let us begin by proving a frequently used result about positive operators.

***10.4.2. Lemma.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be positive, $\xi \in \mathcal{H}$. If $\langle T\xi, \xi\rangle = 0$, then $T\xi = 0$.*

*Proof.* Define a form

$$[\xi, \eta] = \langle T\xi, \xi\rangle, \qquad \xi, \eta \in \mathcal{H}.$$

One can check directly that this form is sesquilinear, and the positivity of $T$ makes it positive. Then the Cauchy–Schwarz inequality (Theorem 4.2.2) applies, giving us

$$|\langle T\xi, \eta\rangle| \leq \langle T\xi, \xi\rangle^{1/2}\langle T\eta, \eta\rangle^{1/2}, \qquad \eta \in \mathcal{H}.$$

Thus $\langle T\xi, \eta\rangle = 0$ for all $\eta$, and so $T\xi = 0$. □

Next we describe some basic techniques that are frequently used.

***10.4.3. Proposition.***

*Let $A, B \in \mathcal{B}(\mathcal{H})$.*

*(i) if $A \geq 0$, then $A \leq \|A\|\, I$;*

*(ii) if $0 \leq A \leq B$ and $T \in \mathcal{B}(\mathcal{H})$, then $0 \leq TAT^* \leq TBT^*$;*

*(iii) if $A \geq 0$ and invertible, then $A^{-1} \geq 0$;*

*(iv) if $A \geq 0$ then $A \leq I$ if and only if $\|A\| \leq 1$;*

*(v) if $c > 0$, then $A \geq cI$ if and only if $A$ is invertible, positive, and $A^{-1} \leq c^{-1} I$.*

*Proof.* Exercise 10.4.4. □

**10.4.2. Square Roots.** When we eventually have the continuous functional calculus available (Corollary 11.2.6), the following result (existence of **square roots**) will be straightforward. For now, we need an ad-hoc proof, based on our manipulations from Remark 7.4.17.

***10.4.4. Proposition.* *(Positive Square Roots)***

*Let $T \in \mathcal{B}(\mathcal{H})$ be positive. Then there exists a unique $S \in \mathcal{B}(\mathcal{H})$, positive, with $S^2 = T$.*

*Proof.* Assume first that $0 \leq T \leq I$, with $\|T\| < 1$. Using Remark 7.4.17, let

$$S = I - \sum_{n=1}^{\infty} \frac{2(2n-2)!}{4^n n!(n-1)!} (I-T)^n. \tag{10.4}$$

The estimates in Remark 7.4.17 show that the series converges in norm, so $S$ exists in $\mathcal{B}(\mathcal{H})$. Though the series converges numerically to the square root when evaluated at numbers, it is not automatically obvious that such is the case with operators. But we can do the following. Again from Remark 7.4.17 we can write

$$S = \sum_{n=0}^{\infty} (-1)^n \binom{1/2}{n} (I-T)^n.$$

Since the series converges uniformly in $\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}$, this allows us to exchange limits and then

$$S^2 = \sum_{n,m=0}^{\infty} (-1)^{n+m} \binom{1/2}{n}\binom{1/2}{m} (I-T)^{n+m}$$

$$= \sum_{m=0}^{\infty} (-1)^m (I-T)^m \sum_{k=0}^{m} \binom{1/2}{k}\binom{1/2}{m-k}$$
$$= I - (I-T) = T$$

by using Lemma 7.4.18. This shows that $S^2 = T$. Also, for any $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ we have

$$\langle (I-T)^n \xi, \xi\rangle \le \|(I-T)^n\| \le \|I-T\|^n \le 1$$

and then, by continuity of the inner product (to exchange the inner product with the series),

$$\langle S\xi, \xi\rangle = 1 - \sum_{n=1}^{\infty} \frac{2(2n-2)!}{4^n n!(n-1)!} \langle (I-T)^n \xi, \xi\rangle \ge 1 - \sum_{n=1}^{\infty} \frac{2(2n-2)!}{4^n n!(n-1)!} = 0.$$

So $S \ge 0$. For arbitrary $T \in \mathcal{B}(\mathcal{H})^+$, do the above for $T/(\|T\|+1)$ and rescale $S$ appropriately at the end.

Now the uniqueness. Things would be easier with functional calculus (Section 11.2), but here is an argument that will do for now. Suppose that $R \ge 0$ and that $T = R^2 = S^2$. We have $RT = R^3 = TR$; and since $S$ is a limit of polynomials on $T$, it follows that $RS = SR$; in particular, $RS$ is selfadjoint. Let $\xi \in \mathcal{H}$ and $\eta = (R-S)\xi$. Then

$$(R+S)\eta = (R+S)(R-S)\xi = (R^2 - S^2)\xi = 0.$$

Also,

$$0 \le \langle R\eta, \eta\rangle \le \langle R\eta, \eta\rangle + \langle S\eta, \eta\rangle = \langle (R+S)\eta, \eta\rangle = 0,$$

so $\langle R\eta, \eta\rangle = 0$. By Lemma 10.4.2, $R\eta = 0$. This we can write as $(R^2 - RS)\xi = 0$. Then

$$\|(R-S)\xi\|^2 = 2\langle R^2\xi, \xi\rangle - 2\mathrm{Re}\,\langle R\xi, S\xi\rangle = 2\langle R^2\xi, \xi\rangle - 2\mathrm{Re}\,\langle \xi, RS\xi\rangle$$
$$= 2\langle \xi, R^2\xi\rangle - 2\langle \xi, RS\xi\rangle = 0.$$

As this occurs for all $\xi \in \mathcal{H}$, $R = S$. □

As is standard, we will denote the $S$ found in Proposition 10.4.4 by $T^{1/2}$. As one simple example of how the square root can be used, here is another proof of Lemma 10.4.2: if $T \ge 0$ and $\langle T\xi, \xi\rangle = 0$, then

$$0 = \langle T\xi, \xi\rangle = \langle T^{1/2}\xi, T^{1/2}\xi\rangle = \|T^{1/2}\xi\|^2.$$

Hence $T^{1/2}\xi = 0$, and multiplying by $T^{1/2}$ we get $T\xi = 0$.

*10.4.5. Remark.* When $T$ is diagonal and positive, its square root is obtained by taking the square roots of the diagonal entries. Other than this and projections (where a projection $P$ is its own square root), there are very cases where finding a square root explicitly is useful or even feasible. Finding the unique positive square root of concrete operators is more of an art than a science. For matrices, things can formally be described in simple terms: if $A \in M_n(\mathbb{C})$ is positive, then $A = UDU^*$ for some unitary $U$ and diagonal $D$ (by Schur triangularization, say). Then $A^{1/2} = UD^{1/2}U^*$, and $D^{1/2}$ can be simply calculated by taking square roots of its entries (since $D$ is diagonal). If $A$ is diagonal, say

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix},$$

then one immediately gets that $A^{1/2} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{5} \end{bmatrix}$. Performing this in any nontrivial example is not fun, though. Consider

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}.$$

This is positive, since its diagonal entries and its determinant are positive; that is, if $\lambda_1, \lambda_2$ are the eigenvalues of $A$, then $\lambda_1 + \lambda_2 = \operatorname{Tr}(A) = 5$, $\lambda_1\lambda_2 = \det A = 2$. This forces $\lambda_1 > 0$, $\lambda_2 > 0$, and so $A$ is positive since it is also selfadjoint. So far so good. Well, the square root of $A$ happens to be

$$A^{1/2} = \sqrt{\frac{31 + 8\sqrt{2}}{17}} \begin{bmatrix} 1 & \frac{6-2\sqrt{2}}{7} \\ \frac{6-2\sqrt{2}}{7} & \frac{4+\sqrt{2}}{7} \end{bmatrix}.$$

The more perceptive readers will probably come to the correct conclusion that going from $A$ to $A^{1/2}$ requires a bit more than a couple of additions and subtractions. First hint that things are complicated comes when one notices that the eigenvalues of $A$ are

$$\lambda_1 = \frac{5 + \sqrt{17}}{2}, \qquad \lambda_2 = \frac{5 - \sqrt{17}}{2}.$$

But comparing this with the expression for $A^{1/2}$ shows that there is still quite a bit of work to be done. In summary, we won't usually be trying to compute square roots explicitly. They are still incredibly useful, though.

The square root is **operator-monotone**:

***10.4.6. Proposition.***

*Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive, with $0 \le S \le T$. Then $S^{1/2} \le T^{1/2}$.*

*Proof.* Assume first that $T$ is invertible. We have

$$T^{-1/2}ST^{-1/2} \le T^{-1/2}TT^{-1/2} = I.$$

By Proposition 10.4.3 we get

$$\|S^{1/2}T^{-1/2}\| = \|T^{-1/2}ST^{-1/2}\|^{1/2} \le 1.$$

The equality

$$T^{-1/4}S^{1/2}T^{-1/4} = T^{-1/4}(S^{1/2}T^{-1/2})\,T^{1/4}$$

is a similarity, so $T^{-1/4}S^{1/2}T^{-1/4}$ and $S^{1/2}T^{-1/2}$ have the same spectral radius. Then, using Proposition 10.3.3,

$$\begin{aligned}\|T^{-1/4}S^{1/2}T^{-1/4}\| &= \operatorname{spr}(T^{-1/4}S^{1/2}T^{-1/4}) = \operatorname{spr}(S^{1/2}T^{-1/2}) \\ &\le \|S^{1/2}T^{-1/2}\| \le 1.\end{aligned}$$

Using Proposition 10.4.3 again, we have

$$T^{-1/4}S^{1/2}T^{-1/4} \le I,$$

which after conjugating with $T^{1/4}$ gives us $S^{1/2} \le T^{1/2}$.

For general $T$, fix $\varepsilon > 0$ and note that $T + \varepsilon I$ is invertible (since $\sigma(T + \varepsilon I) \subset [\varepsilon, \infty)$) and $S \le T + \varepsilon I$. By the above we get that $S^{1/2} \le (T + \varepsilon I)^{1/2}$. So it is enough to show that $(T+\varepsilon I)^{1/2} \to T^{1/2}$. For this, we note that since $T + \varepsilon I$ and $T$ commute we also have that $(T + \varepsilon I)^{1/2}$ and $T^{1/2}$ commute (Exercise 10.4.9). Also, $T + \varepsilon I \ge \varepsilon I$, so

$$(T + \varepsilon I)^{1/2} + T^{1/2} \ge (T + \varepsilon I)^{1/2} \ge \varepsilon^{1/2}\, I,$$

which shows that $(T + \varepsilon I)^{1/2} + T^{1/2}$ is invertible and that

$$\big((T + \varepsilon I)^{1/2} + T^{1/2}\big)^{-1} \le \varepsilon^{-1/2}\, I.$$

Then, from the elementary identity $a^2 - b^2 = (a - b)(a + b)$ for commuting $a, b$,

$$\begin{aligned}\|(T + \varepsilon I)^{1/2} - T^{1/2}\| &= \big\|\big((T + \varepsilon I) - T\big)\big((T + \varepsilon I)^{1/2} + T^{1/2}\big)^{-1}\big\| \\ &= \varepsilon\,\big\|\big((T + \varepsilon I)^{1/2} + T^{1/2}\big)^{-1}\big\| \\ &\le \varepsilon\,\varepsilon^{-1/2} = \varepsilon^{1/2}.\end{aligned}$$

This shows that $(T + \varepsilon I)^{1/2} \xrightarrow[\varepsilon\to 0]{} T^{1/2}$, and thus $S^{1/2} \le T^{1/2}$. $\square$

If the result in Proposition 10.4.6 seems too obvious to deserve a statement and a longish proof on its own, compare with the fact that taking squares is **not** operator-monotone in general (Exercise 10.4.8).

A basic application of square roots is the following characterization of positivity.

***10.4.7. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$. The following statements are equivalent:*

*(i)* $T \ge 0$*;*

*(ii) there exists* $S \in \mathcal{B}(\mathcal{H})$ *with* $S^*S = T$*;*

*(iii)* $T = T^*$ *and* $\sigma(T) \subset [0, \infty)$*.*

*Proof.* (i) $\Longrightarrow$ (ii) This follows directly from Proposition 10.4.4, with $S = T^{1/2}$.

(ii) $\Longrightarrow$ (i) We have, for any $\xi \in \mathcal{H}$,

$$\langle T\xi, \xi\rangle = \langle S^*S\xi, \xi\rangle = \langle S\xi, S\xi\rangle \geq 0.$$

(i) $\Longrightarrow$ (iii) By Proposition 10.3.6 we have that $T = T^*$. And by Proposition 10.2.3, for any $\xi \in \mathcal{H}$ we have

$$\sigma(T) \subset \operatorname{conv}\sigma(T) \subset \overline{W(T)} \subset [0, \infty).$$

(iii) $\Longrightarrow$ (i) Suppose initially that $\|T\| \leq 1$. This guarantees that $\sigma(I - T) \subset [0, 1]$. And, more generally, since $(I - T)^n$ is selfadjoint, by Proposition 10.1.8 we have $\|(I - T)^n\| = \operatorname{spr}\big((I - T)^n\big) \leq 1$. Then the series (10.4) in the proof of Proposition 10.4.4 defines an operator $S$ with $S^2 = T$, with the exact argument as in the first part of the proof of Proposition 10.4.4. If we prove that $S$ is selfadjoint, then $T = S^2 = S^*S$ is positive. And this is the case, for the series in (10.4) is a limit of real linear combinations of selfadjoints, hence selfadjoint (recall that the operation of taking adjoint is isometric, so in particular it is continuous).

For arbitrary $T$, with $c = \frac{1}{2\|T\|}$ the operator $cT$ satisfies $(cT)^* = cT$ and $\sigma(cT) = c\sigma(T) \subset [0, 1]$. Then $cT \geq 0$, which implies that $T \geq 0$. ☐

*10.4.8. Remark* . Given selfadjoint $T$ with $\|T\| \leq 1$, we can define $U = T + i(I - T^2)^{1/2}$. This is well-defined because $\|T\| \leq 1$ guarantees that $I - T^2 \geq 0$ (Proposition 10.4.3 and Exercise 10.4.7). We have

$$\begin{aligned} U^*U &= (T - i(I - T^2)^{1/2})(T + i(I - T^2)^{1/2}) \\ &= T^2 + I - T^2 - i(I - T^2)T + iT(I - T^2) = I. \end{aligned}$$

Similarly $UU^* = I$, and so $U$ is a unitary. Hence $T = \frac{1}{2}\left[T + i(I - T^2)^{1/2}\right] + \frac{1}{2}\left[T - i(I - T^2)^{1/2}\right]$ is a linear combination of unitaries. By scaling we get the same assertion for any selfadjoint $T$. And since arbitrary $T \in \mathcal{B}(\mathcal{H})$ can be written as $T = \operatorname{Re} T + i \operatorname{Im} T$, a linear combination of selfadjoint operators, we have shown that any $T \in \mathcal{B}(\mathcal{H})$ is a linear combination of four unitaries.

### 10.4.3. Partial Isometries and the Polar Decomposition.

**10.4.9. Definition.**

*Let $K, L \subset \mathcal{H}$ be subspaces. An operator $V \in \mathcal{B}(\mathcal{H})$ is a* ***partial isometry*** *with initial space $K$ and final space $L$, if $\ker V = K^\perp$, $\operatorname{ran} V = L$, and $\|V\xi\| = \|\xi\|$ for all $\xi \in K$.*

If $K$ is closed—as it will be usually considered—then $L$ is automatically closed since $V$ is bounded below (Exercise 10.3.5).

***10.4.10. Proposition.***

*Let $V \in \mathcal{B}(\mathcal{H})$, and $K, L \subset \mathcal{H}$ closed subspaces. The following statements are equivalent:*

*(i) $V$ is a partial isometry with initial space $K$ and final space $L$;*

*(ii) $V^*$ is a partial isometry with initial space $L$ and final space $K$;*

*(iii) $V^*V$ is the orthogonal projection onto $K$, and $VK = L$;*

*(iv) $VV^*$ is the orthogonal projection onto $L$, and $V^*L = K$.*

*Proof.* By Polarization, from $\|V\xi\| = \|\xi\|$ we get that $\langle V\xi, V\eta\rangle = \langle \xi, \eta\rangle$ for all $\xi, \eta \in \mathcal{H}$. We will use Proposition 10.1.10 repeatedly.

(i) $\Longrightarrow$ (ii) We have $\ker V^* = (\operatorname{ran} V)^\perp = L^\perp$. And $\overline{\operatorname{ran} V^*} = (\ker V)^\perp = K$. So $V^*$ has initial space $L$ and final space $K$. If $\xi \in L$, we may write $\xi = V\eta$. Then

$$\|V^*\xi\|^2 = \langle V^*\xi, V^*\xi\rangle = \langle V^*V\eta, V^*V\eta\rangle = \langle V\eta, VV^*V\eta\rangle$$
$$= \langle \eta, V^*V\eta\rangle = \langle V\eta, V\eta\rangle = \langle \xi, \xi\rangle = \|\xi\|^2.$$

So $V^*$ is a partial isometry with initial space $L$ and final space $K$.

(ii) $\Longrightarrow$ (i) Here we can repeat the argument with the roles of $V$ and $V^*$ (and $K$ and $L$) exchanged.

(i) $\Longrightarrow$ (iii) We have $\ker V^*V = \ker V = K^\perp$ and

$$\langle V^*V(\xi_K + \xi_{K^\perp}), \eta\rangle = \langle V\xi_K, V\eta\rangle = \langle \xi_K, \eta\rangle.$$

As this can be done for any $\eta$, we have that $V^*V(\xi_K + \xi_{K^\perp}) = \xi_K$. So $V^*V = P_K$. From $\ker V = K^\perp$ and $\operatorname{ran} V = L$ we get $VK = L$.

(iii) $\Longrightarrow$ (i) If we write $\xi = \xi_K + \xi_{K^\perp}$, then

$$\|V\xi\|^2 = \langle V^*V\xi, \xi\rangle = \langle \xi_K, \xi_K\rangle = \|\xi_K\|^2.$$

So, when $\xi = \xi_K \in K$ we get $\|V\xi\| = \|\xi\|$. Since $\ker V = \ker V^*V = K^\perp$, we have that $\operatorname{ran} V = VK = L$.

(ii) $\Longleftrightarrow$ (iv) Apply (ii) $\Longleftrightarrow$ (iii) to $V^*$ with initial space $L$ and final space $K$. $\square$

The **polar decomposition**, below, is the operator analog of the polar form of a complex number. For a closed subspace $K$ we denote by $[K]$ the orthogonal projection onto $K$.

***10.4.11. Proposition. (Polar Decomposition)***

*Let $T \in \mathcal{B}(\mathcal{H})$. Then there exists a unique partial isometry $V$ with $V^*V = [\operatorname{ran} T^*]$, $VV^* = [\operatorname{ran} T]$, and $T = V|T|$.*

*Proof.* Since $T^*T \geq 0$, the operator $|T| = (T^*T)^{1/2}$ exists by Proposition 10.4.4. Define $V$ on $\operatorname{ran}|T|$ by

$$V|T|x = Tx.$$

We need to check that this is well-defined: and it is, since by Proposition 10.1.10 we have $\ker|T| = \ker T$. On $(\operatorname{ran}|T|)^\perp = \ker|T| = \ker T$, we define $V$ as 0. Then

$$\|V|T|x\|^2 = \|Tx\|^2 = \langle T^*Tx, x\rangle = \langle |T|^2 x, x\rangle = \|\,|T|x\,\|^2,$$

so $V$ is isometric on $\operatorname{ran}|T|$ and thus it extends to an isometry on $\overline{\operatorname{ran}}|T|$, with range $\overline{\operatorname{ran}}T$. In particular, $V$ is a partial isometry and so $V^*V$ and $VV^*$ are projections onto $\overline{\operatorname{ran}}|T|$ and $\overline{\operatorname{ran}}T$ respectively.

By definition, $V|T| = T$. Also,

$$\operatorname{ran} V^*V = \overline{\operatorname{ran}}|T| = (\ker|T|)^\perp = (\ker T)^\perp = \overline{\operatorname{ran}}T^*.$$

For the uniqueness, suppose that $T = W|T|$ and $W^*W$ is the orthogonal projection onto $\overline{\operatorname{ran}}T^*$. Then $\ker W = \ker W^*W = (\operatorname{ran} T^*)^\perp = \ker V^*V = \ker V$. And on $\operatorname{ran} V^*V = \overline{\operatorname{ran}}T^* = \overline{\operatorname{ran}}|T|$ we have $W|T|x = Tx = V|T|x$, so the equality extends to the closure and $W = V$. $\square$

The uniqueness of the polar decomposition is actually a bit stronger: if $T = WZ = VS$ with $W, V$ partial isometries such that $V^*V = \overline{\operatorname{ran}}Z = \overline{\operatorname{ran}}V = W^*W$, $VV^* = WW^* = \overline{\operatorname{ran}}T$, and $Z, S \geq 0$, then $W = V$ and $Z = S$ (Exercise 10.4.14). It is important to emphasize, though, that the uniqueness statement requires the condition on the ranges.

**10.4.4. Operator-Block Form.** We saw on (9.1.2), page 552, that given a bounded projection from a Banach space into a closed subspace it is possible to see operators as operator-block matrices. Of course this works in the case of Hilbert spaces. As closed subspaces are always complemented in a Hilbert space, via a canonical orthogonal projection, we will only consider those decompositions. We have

**10.4.12. Proposition.**

*Let $\mathcal{H}_1 \subset \mathcal{H}$ be a closed subspace and $T \in \mathcal{B}(\mathcal{H})$. Then*

$$T^* = \begin{bmatrix} T_{11}^* & T_{21}^* \\ T_{12}^* & T_{22}^* \end{bmatrix}$$

*Proof.* Exercise 10.4.22. □

**10.4.13. Corollary.**

*Let $T \in \mathcal{B}(\mathcal{H})$ be positive and have, with respect to a certain closed subspace $\mathcal{H}_1 \subset \mathcal{H}$, the operator-block form*

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & 0 \end{bmatrix}.$$

*Then $T_{12} = 0$, $T_{21} = 0$.*

*Proof.* We know from Proposition 10.4.7 that there exists $S \in \mathcal{B}(\mathcal{H})$ with $T = S^*S$. Then, knowing the form of $S^*$ from Proposition 10.4.12,

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & 0 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}^* \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$
$$= \begin{bmatrix} S_{11}^*S_{11} + S_{21}^*S_{21} & S_{11}^*S_{12} + S_{21}^*S_{22} \\ S_{21}^*S_{11} + S_{22}^*S_{21} & S_{12}^*S_{12} + S_{22}^*S_{22} \end{bmatrix}.$$

Looking at the $2,2$ entry we have $S_{12}^*S_{12} + S_{22}^*S_{22} = 0$, which immediately gives $S_{21} = S_{22} = 0$ (via Exercise 10.4.2). Thus $T_{12} = 0$, $T_{21} = 0$. □

We can also use block matrices in a kind of reverse way, in the sense that starting from $\mathcal{B}(\mathcal{H})$ we can see $\mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ as the $2 \times 2$ matrices on $\mathcal{B}(\mathcal{H})$, through the decomposition $\mathcal{H} \oplus \mathcal{H} = (\mathcal{H} \oplus 0) \oplus (0 \oplus \mathcal{H})$.

It is possible to see positivity of an operator from its operator-block matrix form, but the result is not particularly pretty (Corollary 10.4.15). But some special cases are decently nice:

**10.4.14. Proposition.**

*Let $S, T \in \mathcal{B}(\mathcal{H})$. Then $\begin{bmatrix} I & S \\ S^* & T \end{bmatrix} \geq 0$ if and only if $S^*S \leq T$.*

*Proof.* The operator $\begin{bmatrix} I & -S \\ 0 & I \end{bmatrix}$ is invertible with inverse $\begin{bmatrix} I & S \\ 0 & I \end{bmatrix}$. Using Exercise 10.4.5,

$$\begin{aligned}\begin{bmatrix} I & S \\ S^* & T \end{bmatrix} \geq 0 &\iff \begin{bmatrix} I & 0 \\ -S^* & I \end{bmatrix} \begin{bmatrix} I & S \\ S^* & T \end{bmatrix} \begin{bmatrix} I & -S \\ 0 & I \end{bmatrix} \geq 0 \\ &\iff \begin{bmatrix} I & 0 \\ 0 & T - S^*S \end{bmatrix} \geq 0 \iff T - S^*S \geq 0.\end{aligned}$$

□

**10.4.15. Corollary.**

*Let $R, S, T \in \mathcal{B}(\mathcal{H})$. The following statements are equivalent:*

*(i)* $\begin{bmatrix} R & S \\ S^* & T \end{bmatrix} \geq 0$*;*

*(ii)* $R, T \geq 0$ *and for each* $\varepsilon > 0$ *there exists* $X_\varepsilon \in \mathcal{B}(\mathcal{H})$*, contractive, with* $S = (R + \varepsilon I)^{1/2} X_\varepsilon (T + \varepsilon I)^{1/2}$.

*Proof.* (i) $\implies$ (ii) We have

$$\langle R\xi, \xi \rangle = \left\langle \begin{bmatrix} R & S \\ S^* & T \end{bmatrix} \begin{bmatrix} \xi \\ 0 \end{bmatrix}, \begin{bmatrix} \xi \\ 0 \end{bmatrix} \right\rangle \geq 0.$$

So $R \geq 0$ and, similarly, $T \geq 0$. Let $X_\varepsilon = (R+\varepsilon I)^{-1/2} X_\varepsilon (T+\varepsilon I)^{-1/2}$. Then, with $R_\varepsilon = (R + \varepsilon I)^{-1/2}$ and $T_\varepsilon = (T + \varepsilon I)^{-1/2}$,

$$\begin{bmatrix} I & X_\varepsilon \\ X_\varepsilon^* & I \end{bmatrix} = \begin{bmatrix} R_\varepsilon & 0 \\ 0 & T_\varepsilon \end{bmatrix} \begin{bmatrix} R + \varepsilon I & S \\ S^* & T + \varepsilon I \end{bmatrix} \begin{bmatrix} R_\varepsilon & 0 \\ 0 & T_\varepsilon \end{bmatrix} \geq 0.$$

By Proposition 10.4.14 this gives $X_\varepsilon^* X_\varepsilon \leq I$, and so $\|X_\varepsilon\| \leq 1$ with $S = (R + \varepsilon I)^{1/2} X_\varepsilon (T + \varepsilon I)^{1/2}$.

(ii) $\implies$ (i) Fix $\varepsilon > 0$. We can write

$$\begin{aligned}\begin{bmatrix} R + \varepsilon I & S \\ S^* & T + \varepsilon I \end{bmatrix} &= \begin{bmatrix} R + \varepsilon I & (R + \varepsilon I)^{1/2} X_\varepsilon (T + \varepsilon I)^{1/2} \\ (T + \varepsilon I)^{1/2} X_\varepsilon^* (R + \varepsilon I)^{1/2} & T + \varepsilon I \end{bmatrix} \\ &= Y \begin{bmatrix} I & X_\varepsilon \\ X_\varepsilon^* & I \end{bmatrix} Y^*,\end{aligned}$$

where

$$Y = \begin{bmatrix} (R + \varepsilon I)^{1/2} & 0 \\ 0 & (T + \varepsilon I)^{1/2} \end{bmatrix}.$$

Since $\|X_\varepsilon\| \leq 1$, we have $\begin{bmatrix} I & X_\varepsilon \\ X_\varepsilon^* & I \end{bmatrix} \geq 0$ by Proposition 10.4.14, and thus

$$\begin{bmatrix} R + \varepsilon I & S \\ S^* & T + \varepsilon I \end{bmatrix} \geq 0$$

for all $\varepsilon > 0$, which gives (i). □

**10.4.5. Exercises.**

(10.4.1) Show that if $T : \mathcal{H} \to \mathcal{H}$ is linear and $\langle T\xi, \xi\rangle \geq 0$ for all $\xi \in \mathcal{H}$, then $T$ is bounded.

(10.4.2) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive with $S + T = 0$. Show that $S = T = 0$.

(10.4.3) Let $A \in \mathcal{B}(\mathcal{H})$ be positive and invertible. Show that $A^{1/2}$ is invertible.

(10.4.4) Prove Proposition 10.4.3.

(10.4.5) Let $T, S \in \mathcal{B}(\mathcal{H})$, with $S$ invertible. Show that $T \geq 0$ if and only if $S^*TS \geq 0$.

(10.4.6) Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. Show that $T^*$ exists by using block matrices over $\mathcal{H} \oplus \mathcal{K}$ and Theorem 10.1.6. For this, consider $X \in \mathcal{B}(\mathcal{H} \oplus \mathcal{K})$ with $X_{21} = T$ and $X_{11} = 0$, $X_{12} = 0$, $X_{22} = 0$.

(10.4.7) Let $T \in \mathcal{B}(\mathcal{H})$ with $0 \leq T \leq I$. Show that $T^2 \leq T$.

(10.4.8) Show an example of positive $S, T \in \mathcal{B}(\mathcal{H})$ such that $S \leq T$ but $S^2 \not\leq T^2$. *(Hint: examples already exist on dimension* 2*)*

(10.4.9) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive and such that $ST = TS$. Show that $S^{1/2}T^{1/2} = T^{1/2}S^{1/2}$.

(10.4.10) Show an example of $T \in \mathcal{B}(\mathcal{H})$ such that $\operatorname{ran} T$ is not closed.

(10.4.11) Let $T \in \mathcal{B}(\mathcal{H})$. Show that $\|T\xi\| = \|\,|T|\xi\|$ for all $\xi \in \mathcal{H}$, and that $|T|$ is the only positive operator in $\mathcal{B}(\mathcal{H})$ with that property.

(10.4.12) Let $V \in \mathcal{B}(\mathcal{H})$ be a partial isometry with $P = V^*V$ and $Q = VV^*$. Show that $V = QVP$.

(10.4.13) Show that when $\dim \mathcal{H} < \infty$ the partial isometry in Proposition 10.4.11 can be chosen to be a unitary, at the cost of losing the condition on the range.

(10.4.14) Show that if $T = WZ = VS$ with $W, V$ partial isometries with $V^*V = [\overline{\operatorname{ran}} Z] = [\overline{\operatorname{ran}} S] = W^*W$, $VV^* = WW^* = [\overline{\operatorname{ran}} T]$, and $Z, S \geq 0$, then $W = V$ and $Z = S$.

(10.4.15) Let $T \in \mathcal{B}(\mathcal{H})$ be invertible. Show that if $T = V|T|$ is the Polar Decomposition, then $|T|$ is invertible and $V$ is a unitary.

(10.4.16) Give an example of $T \in \mathcal{B}(\mathcal{H})$ and decompositions $T = VR = WS$, with $R, S \geq 0$ and $V, W$ partial isometries, such that $R \neq S$ and $V \neq W$. Explain why this does not contradict the uniqueness in Proposition 10.4.11.

(10.4.17) Let $T \in \mathcal{B}(\mathcal{H})$. Show $T$ can be expressed as $T = WA$ with $A$ positive and $W$ a unitary if and only if $\dim\ker T = \dim\ker T^*$.

(10.4.18) Let $T \in \mathcal{B}(\mathcal{H})$, invertible. Show that $\|T\| = \|T^{-1}\| = 1$ if and only if $T$ is a unitary.

(10.4.19) Let $S, T \in \mathcal{B}(\mathcal{H})$ be positive. Show that $\sigma(ST) \subset [0, \infty)$. Does this imply that $ST$ is positive?

(10.4.20) Let $T \in \mathcal{B}(\mathcal{H})^+$ with $\|T\| \leq 1$, and $\xi_0 \in \mathcal{H}$ such that $\|T\xi_0\| = \|\xi_0\|$. Show that $T\xi_0 = \xi_0$, and that $T = P + T_0$, where $P$ is a projection and $T_0 \in \mathcal{B}(\mathcal{H})$ satisfies $\|T_0\xi\| < \|\xi\|$ for all nonzero $\xi \in \mathcal{H}$.

(10.4.21) Let $S, T \in \mathcal{B}(\mathcal{H})$, both positive, and with $\|S\| \leq 1$ and $\|T\| \leq 1$. Show that $\|S - T\| \leq 1$.

(10.4.22) Prove Proposition 10.4.12.

## 10.5. Projections

Projections, with a direct relation to closed subspaces, play a role in studying Hilbert spaces, their operators, and particularly von Neumann algebras. The do also play a central role in C*-theory, but such role is beyond the scope of this book. We mention a few basic facts here. A lot more about projections will be said in Chapter 14, particularly in Section 14.2.

**10.5.1. Definition.**

*A **projection** (or, an **orthogonal projection**) is $P \in \mathcal{B}(\mathcal{H})$ with $P^*P = P$ (equivalently, $P^* = P$, $P^2 = P$). Given two projections $P, Q$, we say that $Q$ is a **subprojection** of $P$ if $Q\mathcal{H} \subset P\mathcal{H}$.*

It is straightforward from the definition that a projection is always positive. We showed in Proposition 4.3.8 that there is a projection onto any closed subspace of $\mathcal{H}$. Some authors use "orthogonal projection" for what we just named "projection". In such cases they use "projection" for any $P$ with $P^2 = P$ (not necessarily selfadjoint); we will use the term **idempotent** for such $P$.

From $P = P^*P$ it follows that projections are positive operators.

***10.5.2. Proposition.***

*Let $P \in \mathcal{B}(\mathcal{H})$ be nonzero, with $P^2 = P$. Then $\|P\| \geq 1$, and the following statements are equivalent:*

*(i)* $\|P\| = 1$;

*(ii)* $P \geq 0$;

*(iii)* $P^* = P$.

*Proof.* We have $\|P\| = \|P^2\| \leq \|P\|^2$; so, if $P \neq 0$, then $1 \leq \|P\|$.

(i) $\Longrightarrow$ (ii) For $\xi \in \mathcal{H}$,

$$\|P\xi\|^2 = \langle P\xi, P\xi\rangle = \langle P\xi, P^2\xi\rangle = \langle P^*P\xi, P\xi\rangle \leq \|P^*P\xi\| \, \|P\xi\| \leq \|P\xi\|^2.$$

Therefore we get equality in the Cauchy–Schwarz inequality (Theorem 4.2.2) and so there exists $\lambda \in \mathbb{C}$ with $P\xi = \lambda P^*P\xi$. Looking at the above equalities, if $P\xi \neq 0$, we get $\lambda = 1$. Then $P^*P\xi = P\xi$ for all $\xi$ and $P = P^*P \geq 0$.

(ii) $\Longrightarrow$ (iii) If $P \geq 0$, then $P^* = P$ by Proposition 10.3.6.

(iii) $\Longrightarrow$ (i) We have $\|P\| = \|P^*P\| = \|P\|^2$, so $\|P\| = 1$ since $P \neq 0$. ☐

A frequently used consequence of (10.5.2) is that any orthogonal projection $P$ satisfies $0 \leq P \leq I$.

Subprojections can be readily characterized.

***10.5.3. Proposition.***

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. The following statements are equivalent:*

*(i)* $Q$ *is a subprojection of* $P$;

*(ii)* $PQ = Q$;

*(iii)* $QP = Q$;

*(iv)* $Q \leq P$.

*Proof.* (i) $\Longrightarrow$ (ii) Since $Q\mathcal{H} \subset P\mathcal{H}$, for any $\xi \in \mathcal{H}$ we have $PQ\xi = Q\xi$. Thus $PQ = Q$.

(ii) $\Longrightarrow$ (iii) Taking adjoints in $PQ = Q$ we get $QP = Q$.

(iii) $\Longrightarrow$ (iv) We have $Q = QP = PQ$, so $P - Q = P - PQ = P - PQP = P(I-Q)P \geq 0$.

(i) $\Longrightarrow$ (i) From $P - Q \geq 0$ we get

$$0 \leq (I-P)(P-Q)(I-P) = -(I-P)Q(I-P).$$

As $(I-P)Q(I-P) \geq 0$, it follows that $(I-P)Q(I-P) = 0$. Then $0 = [Q(I-P)]^*Q(I-P)$, so $Q(I-P) = 0$ and $(I-P)Q = 0$. Then for any $\xi \in \mathcal{H}$ we have $Q\xi = PQ\xi + (I-P)Q\xi = PQ\xi \in P\mathcal{H}$. ☐

**10.5.4. Definition.**

*We say that projections $\{P_j\}_j \subset \mathcal{B}(\mathcal{H})$ are* **pairwise orthogonal** *if $P_k P_j = 0$ whenever $k \neq j$. This is equivalent to $P_j\mathcal{H} \perp P_k\mathcal{H}$.*

***10.5.5. Proposition.***

*If projections $P_1, \ldots, P_m \subset \mathcal{B}(\mathcal{H})$ satisfy $\sum_j P_j \leq I$, then they are pairwise orthogonal. In particular, if a sum of projections is a projection, then the projections are pairwise orthogonal.*

*Proof.* Fix $s, t$ with $s \neq t$. From $\displaystyle\sum_j P_j \leq I$, we get $\displaystyle\sum_{j \neq s} P_j \leq I - P_s$. By conjugating with $P_s$,

$$0 \leq P_s\Big(\sum_{j \neq s} P_j\Big)P_s \leq P_s(I - P_s)P_s = 0.$$

Thus $P_s\Big(\displaystyle\sum_{j \neq s} P_j\Big)P_s = 0$. In particular,

$$0 \leq P_s P_t P_s \leq P_s\Big(\sum_{j \neq s} P_j\Big)P_s = 0.$$

This gives us $P_s P_t P_s = 0$. We may write this as $(P_t P_s)^* P_t P_s = 0$, and so $P_t P_s = 0$. Taking adjoints we get $P_s P_t = 0$. ☐

The result in Proposition 10.5.5 does not hold for general idempotents. Concretely, it was proven in [BES94] that there exist nonzero idempotents $P_1, \ldots, P_5$ in $\mathcal{B}(\ell^2(\mathbb{N}))$ with $P_1 + \cdots + P_5 = I$.

We owe the following argument to Bruce Blackadar. The mention of C$^*$-algebras—still to be defined—does not affect the flow of the text and can be ignored.

***10.5.6. Proposition.***

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. If $\|P - Q\| < 1$, then there exists a unitary $U \in \mathcal{B}(\mathcal{H})$ such that $Q = UPU^*$. If $P, Q \in \mathcal{A}$ for some $C^*$-algebra $\mathcal{A}$, then $U$ can be chosen to be in $\mathcal{A}$.*

*Proof.* Let $X = PQ + (I - P)(I - Q)$. Then

$$I - X = I - PQ - (I + PQ - P - Q) = P + Q - 2PQ$$

$$= P - (2P - I)Q = (2P - I)(P - Q).$$

Using that $2P - I$ is a unitary,

$$\|I - X\| = \|(2P - I)(P - Q)\| = \|P - Q\| < 1.$$

By Proposition 6.2.3, the operator $X$ is invertible. Let $X = U|X|$ be the polar decomposition. We have $|X| = (X^*X)^{1/2} \in \mathcal{A}$, and $U = X|X|^{-1} \in \mathcal{A}$. Also,

$$PX = PQ = XQ, \quad QX^*X = QPQ.$$

Taking adjoints, $QX^*X = X^*XQ$, and then $Q|X| = |X|Q$ by Exercise 10.4.9. Thus

$$PU = PX|X|^{-1} = XQ|X|^{-1} = U|X|Q|X|^{-1} = UQ|X|\,|X|^{-1} = UQ.$$

As $U$ is a unitary, $Q = U^*PU$. □

*10.5.7. Remark*. Another approach for Proposition 10.5.6, following ideas of Dixmier [Dix48] and Halmos [Hal69], consists of first reducing the problem to

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} C^2 & CS \\ CS & S^2 \end{bmatrix}, \quad C^2 + S^2 = I$$

where $C$ and $S$ are commuting positive contractions; this is inspired by the rank-2 case where $C$ and $S$ would be the cosine and sine (see Proposition 10.5.12 for details). Then one takes

$$U = \begin{bmatrix} C & S \\ S & -C \end{bmatrix}$$

and shows that $UPU^* = Q$ with $U = (P + Q - I)\,|P + Q - I|^{-1} \in \mathcal{A}$.

The reduction mentioned in Remark 10.5.7 and that will be considered in Propositions 14.3.6 and 10.5.12 requires a few notions about projections that will also be useful later.

**10.5.8. Definition.**

*Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ be projections. The* ***intersection*** *of the family $P_j$ is the projection $\bigwedge_j P_j$ defined as the orthogonal projection onto $\bigcap_j P_j\mathcal{H}$.*

*The* ***union*** *of the family $\{P_j\}$ is the projection $\bigvee_j P_j$ defined as the orthogonal projection onto $\overline{\operatorname{span}}^{\|\cdot\|} \{\bigcup_j P_j\mathcal{H}\}$.*

*When only two (or, finitely many) projections are involved we will use the notation $P \wedge Q$ and $P \vee Q$ respectively.*

*For a projection $P$ we will denote the projection $I_{\mathcal{H}} - P$ by $P^{\perp}$.*

*The projections $P$ and $Q$ are in* ***generic position*** *when*

$$P \wedge Q = P^\perp \wedge Q = P \wedge Q^\perp = P^\perp \wedge Q^\perp = 0.$$

*Given $T \in \mathcal{B}(\mathcal{H})$, its* ***range projection*** *is the orthogonal projection onto $\overline{\operatorname{ran}} T$.*

We note that by definition we have $P \wedge Q \leq P$ and $P \wedge Q \leq Q$.

***10.5.9. Proposition.***

*Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ be a family of projections. Then*

$$\Big(\bigwedge_j P_j\Big)^\perp = \bigvee_j P_j^\perp, \qquad \Big(\bigvee_j P_j\Big)^\perp = \bigwedge_j P_j^\perp$$

*Proof.* Exercise 10.5.7. □

We will need the following lemma right away.

***10.5.10. Lemma.***

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections and let $T = P(I - Q)$. Then the range projection of $T$ is $P - P \wedge Q$.*

*Proof.* We have

$$\ker T^* = \{\eta : \ (I - Q)P\eta = 0\} = \{\eta : \ P\eta = QP\eta\}.$$

Let $L$ be the orthogonal projection onto the $\ker T^*$. Since $T^*(I - P) = 0$, we have $(I - P)\mathcal{H} \subset \ker T^*$, and so $I - P \leq L$. We also have $T^*(P \wedge Q) = (I - Q)P(P \wedge Q) = (I - Q)(P \wedge Q) = 0$, so $P \wedge Q \leq L$. It follows, since $I - P$ and $P \wedge Q$ have orthogonal ranges, that $I - P + P \wedge Q \leq L$. Let $S = L - (I - P)$. We have $S = SP + S(I - P) = SP$, so $S \leq P$. Also, since $S \leq L$, $T^*S = T^*LS = 0$, so $0 = (I - Q)PS = (I - Q)S$; then $S = QS$ and $S \leq Q$. As $S \leq P$ and $S \leq Q$, $S \leq P \wedge Q$. Then

$$I - P + P \wedge Q \leq L \leq I - P + P \wedge Q,$$

giving us $L = I - P + P \wedge Q = I - (P - P \wedge Q)$. The projection onto $\overline{\operatorname{ran}} T = (\ker T^*)^\perp$ is then $I - L = P - P \wedge Q$. □

The next result is better phrased in terms of equivalence of projections (Section 14.2) but we are in position to state it and prove it now.

**10.5.11. Proposition. (Kaplansky's Formula)**

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Then there exists a partial isometry $V \in \mathcal{B}(\mathcal{H})$ with $V^*V = P \vee Q - Q$ and $VV^* = P - P \wedge Q$.*

*Proof.* Let $T = P(I - Q)$. By Lemma 10.5.10, the range projection for $T$ is $P - P \wedge Q$. Applying Lemma 10.5.10 to $T^* = (I - Q)P$, we get that its range projection is $I - Q - (I - Q) \wedge (I - P)$. Using Proposition 10.5.9, the range projection of $T^*$ can be rewritten as

$$I - Q - (I - Q) \wedge (I - P) = -Q + \big(I - (I - Q) \wedge (I - P)\big) = -Q + Q \vee P = P \vee Q - Q.$$

The Polar Decomposition (Proposition 10.4.11) then provides us with a partial isometry $V$ with $V^*V = P \vee Q - Q$ and $VV^* = P - P \wedge Q$. □

We will eventually see several applications of Kaplansky's formula. For now let us get the following result, which also shows operator-block matrix manipulations.

**10.5.12. Proposition.**

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be in generic position. Then there exists a unitary $W : \mathcal{H} \to P\mathcal{H} \oplus P\mathcal{H}$ such that*

$$WPW^* = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad WQW^* = \begin{bmatrix} C^2 & CS \\ CS & S^2 \end{bmatrix}, \quad C^2 + S^2 = I, \tag{10.5}$$

*where $C, S \in \mathcal{B}(P\mathcal{H})$ are commuting positive contractions.*

*Proof.* By hypothesis,

$$P \wedge Q = P^\perp \wedge Q = P \wedge Q^\perp = P^\perp \wedge Q^\perp = 0. \tag{10.6}$$

We claim that $B_0 = PQ(I - P)$ is injective on $(I - P)\mathcal{H}$, with dense range in $p\mathcal{H}$. Indeed, if $PQ(I - P)\xi = 0$, then $Q(I - P)\xi \in \operatorname{ran} Q \cap \ker P = \operatorname{ran} Q \cap \operatorname{ran}(I - P) = \operatorname{ran}(Q \wedge (I - P)) = \{0\}$. Then $Q(I - P)\xi = 0$, which means that $(I - P)\xi \in \ker Q \cap \operatorname{ran}(I - P) = \operatorname{ran}(Q^\perp \wedge P^\perp) = \{0\}$. Hence $(I - P)\xi = 0$ and $B_0$ is injective, and $\operatorname{ran} B_0^*$ is dense. We can see this assertion about the range by noticing that $B_0 : (I - P)\mathcal{H} \to P\mathcal{H}$ and

$$\langle B_0(I - P)\xi, P\eta\rangle = \langle (I - P)\xi, (I - P)QP\eta\rangle,$$

showing that $B_0^* = (I - P)QP$ (this needs to done like this and not by "simply applying $*$" because we are thinking of adjoints on spaces other than the original $\mathcal{H}$, and in general changing the codomain changes the adjoint); this $B_0^*$ has $\operatorname{ran} B_0^* = (\ker B_0)^\perp = \{0\}^\perp = (I - P)\mathcal{H}$. Let $B_0 = UH$ be the polar decomposition. Then $U$ is an isometry from $\overline{\operatorname{ran}} B_0^* = (I-P)\mathcal{H}$ onto $\overline{\operatorname{ran}} B_0 = P\mathcal{H}$. That is, $U$ is a unitary $U : (I-P)\mathcal{H} \to P\mathcal{H}$.

Let $W : \mathcal{H} \to P\mathcal{H} \oplus P\mathcal{H}$ be given by

$$W\xi = P\xi \oplus U(I-P)\xi.$$

This $W$ is a symmetry, for $\|W\xi\|^2 = \|P\xi\|^2 + \|U(I-P)\xi\|^2 = \|P\xi\|^2 + \|(I-P)\xi\|^2 = \|\xi\|^2$. And it is surjective, for $P\xi \oplus P\eta = W(P\xi + U^*P\eta)$. Thus $W$ is a unitary. From this last formula and $W^* = W^{-1}$ we get $W^*(P\xi \oplus P\eta) = P\xi + U^*P\eta$. Let $\tilde{P} = WPW^*$. We have

$$\tilde{P}(P\xi \oplus P\eta) = WP(P\xi + (I-P)U^*P\eta) = WP\xi = P\xi \oplus 0.$$

This shows that $\tilde{P} = WPW^*$ is the orthogonal projection onto $P\mathcal{H} \oplus 0$, which is precisely $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$. So we can focus on the form of $WQW^*$. Let $\tilde{Q} = WQW^*$. We can assume that

$$\tilde{Q} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix},$$

where the $2,1$ entry is a consequence of $WQW^*$ being selfadjoint. From $\tilde{Q}^2 = \tilde{Q}$,

$$\begin{aligned} A^2 + BB^* &= A, & AB + BD &= B, \\ B^*A + DB^* &= B^*, & B^*B + D^2 &= D. \end{aligned} \tag{10.7}$$

We have

$$\begin{aligned} W(PU(I-P))W^*(P\xi \oplus P\eta) &= WPU(I-P)(P\xi + U^*P\eta) \\ &= WPUU^*P\eta = WP\eta = P\eta \oplus 0. \end{aligned}$$

That is,

$$W\big(PU(I-P)\big)W^* = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}.$$

We also have

$$\begin{aligned} WB_0W^*(P\xi \oplus P\eta) &= WPQ(I-P)(P\xi + U^*P\eta) = WPQU^*P\eta \\ &= PQU^*P\eta \oplus 0 \\ &= B_0U^*P\eta \oplus 0 = (UHU^*)P\eta \oplus 0. \end{aligned}$$

That is,

$$WB_0W^* = \begin{bmatrix} 0 & UHU^* \\ 0 & 0 \end{bmatrix}.$$

Then

$$\begin{aligned} \begin{bmatrix} 0 & B \\ 0 & 0 \end{bmatrix} &= \tilde{P}\tilde{Q}(I_2 - \tilde{P}) = WPW^*WQW^*W(I-P)W^* = W(PQ(I-P)W^* \\ &= WB_0W^* = \begin{bmatrix} 0 & UHU^* \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

We can then conclude that $B \geq 0$. This allows us to rewrite (10.7) as

$$A^2 + B^2 = A, \qquad AB + BD = B, \qquad BA + DB = B, \qquad B^2 + D^2 = D. \tag{10.8}$$

From $\|\tilde{Q}\| \leq 1$ we know that $\|A\| \leq 1$, so $0 \leq A \leq I$ and, similarly, $0 \leq D \leq I$. Then $B = (A - A^2)^{1/2} = (D - D^2)^{1/2}$. From this we get that $AB = BA$ and $DB = BD$. This gives us that equality $B(A + D) = B$, which we may write as $B(A + D - I) = 0$. But $B$ is injective, since $B_0$ is; indeed, if $BP\eta = 0$ then $0 = WB_0W^*(0 \oplus P\eta) = WB_0U^*P\eta$, so $B_0U^*P\eta$ and then $U^*P\eta = 0$ since $B_0$ is injective, and then $P\eta = 0$ since $U^*$ is isometric on $P\mathcal{H}$. Hence $A + D = I$. Let $C = A^{1/2}$ and $S = (I - A)^{1/2}$. Then $B = (A - A^2)^{1/2} = \big(A(I - A)\big)^{1/2} = CS$. And $CS = SC$ for both are expressions on $A$. We have therefore established (10.5). □

### 10.5.1. Exercises.

(10.5.1) Let $P \in \mathcal{B}(\mathcal{H})$. Show that $P^*P = P$ if and only if $P = P^* = P^2$.

(10.5.2) Let $P \in \mathcal{B}(\mathcal{H})$ be a projection. Show that $\ker P = P^{\perp}\mathcal{H}$ and that $P^{\perp}$ is the orthogonal projection onto $(P\mathcal{H})^{\perp}$.

(10.5.3) When $\mathcal{H} = \mathbb{C}^2$, we can identify $\mathcal{B}(\mathcal{H})$ with $M_2(\mathbb{C})$. Find all orthogonal projections and all idempotents.

(10.5.4) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, and $T \neq 0, \neq I$. Show that $T$ is a projection if and only if $\sigma(T) = \{0, 1\}$.

(10.5.5) Show that the projections are extreme points in the set $\mathcal{B}(\mathcal{H})_1^+$ of positive operators with norm at most 1 (it is actually true that the projections are the only extreme points, but that will have to wait until Exercise 12.4.4).

(10.5.6) Show that the converse of Proposition 10.5.6 is false. That is, find projections $P, Q \in \mathcal{B}(\mathcal{H})$, unitarily equivalent, and such that $\|P - Q\| = 1$.

(10.5.7) Prove Proposition 10.5.9.

(10.5.8) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Is it true that $P = P \wedge Q + P \wedge Q^{\perp}$? Provide either a proof of a counterexample.

(10.5.9) Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ a family of projections. Show that

$$U\Big(\bigvee_j P_j\Big)U^* = \bigvee_j UP_jU^*.$$

(10.5.10) Let $\mathcal{H}$ be a Hilbert space and $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ an increasing net of projections such that $\bigvee_j P_j = I_{\mathcal{H}}$. Show that for all $\xi \in \mathcal{H}$ we have $\lim_j P_j\xi = \xi$.

(10.5.11) Let $P, R \in \mathcal{B}(\mathcal{H})$ be projections. Show that

$$P \vee R + (I_{\mathcal{H}} - P) \wedge (I_{\mathcal{H}} - R) = I_{\mathcal{H}}.$$

(10.5.12) Let $\mathcal{H}$ be a Hilbert space. Show that there exists an increasing net $\{P_j\}$ of finite-rank projections with $\bigvee_j P_j = I_{\mathcal{H}}$.

(10.5.13) Let $\mathcal{H}$ be a Hilbert space and $\{P_j\}$ an increasing net of projections with $P_j\xi \to \xi$ for all $\xi \in \mathcal{H}$. Show that for all $T \in \mathcal{B}(\mathcal{H})$,

$$\|T\| = \sup\{\|P_jTP_j\| : \ j\}.$$

(10.5.14) Let $T, P \in \mathcal{B}(\mathcal{H})$ with $P$ a projection such that $(I_{\mathcal{H}}-P)T(I_{\mathcal{H}}-P) = 0$. Show that $|\langle T\xi, \xi\rangle| \leq 2\|T\|\,\|P\xi\|$ for all $\xi \in \mathcal{H}$ with $\|\xi\| \leq 1$.

(10.5.15) Let $T \in \mathcal{B}(\mathcal{H})^+$, $\alpha, \beta, \gamma \in \mathbb{R}$ with $0 < \gamma < \alpha < \beta$, and $P, Q \in \mathcal{B}(\mathcal{H})$ two projections that commute with $T$ and with each other, and such that

$$\alpha P \leq PT \leq \beta P, \qquad 0 \leq (I_{\mathcal{H}} - P)T \leq \gamma(I_{\mathcal{H}} - P),$$
$$\alpha Q \leq QT \leq \beta Q, \qquad 0 \leq (I_{\mathcal{H}} - Q)T \leq \gamma(I_{\mathcal{H}} - Q).$$

Show that $P = Q$.

(10.5.16) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections.

(i) Let $T = QP + (I - Q)(I - P)$. Show that $TP = QT$.

(ii) Show that if $\|P - Q\| < 1$, then $T$ is invertible.

(iii) Show that if $\|P - Q\| < 1$ then

$$\dim\operatorname{ran} P = \dim\operatorname{ran} Q,\ \ \dim\operatorname{ran}(I - P) = \dim\operatorname{ran}(I - Q).$$

(iv) Conclude that there exists a unitary $U$ with $UPU^* = Q$.

This proves the result of Proposition 10.5.6, but without writing $U$ as an expression depending on $P$ and $Q$.

(10.5.17) Regarding Remark 10.5.7, show that $U = (P + Q - I)|P + Q - I|^{-1}$ is a unitary and that $UPU^* = Q$.

## 10.6. Compact operators

We have already discussed compact operators on a Banach space in Section 9.6. As Hilbert spaces are Banach spaces, everything said there is valid for compact operators in $\mathcal{B}(\mathcal{H})$. At the same time, there are enough particulars in the case of compact operators on a Hilbert space that this section makes sense. Most notably, as opposed to operators on arbitrary Banach spaces, on a Hilbert space compact operators are always limits of finite-rank operators. And we always have orthonormal bases available, which allow for more precise descriptions.

**10.6.1. Finite-rank operators.** We will denote the set of all $T \in \mathcal{B}(\mathcal{H})$ with finite rank by $\mathcal{F}(\mathcal{H})$. The prescription that $T \in \mathcal{B}(\mathcal{H})$ is necessary, for unbounded linear maps with finite rank exist. For instance fix $\eta \in \mathcal{H}$, take $f : \mathcal{H} \to \mathbb{C}$ unbounded, and put $T\xi = f(\xi)\,\eta$.

Finite-rank operators over a Hilbert space can be represented in a very concrete way. Indeed, given $T \in \mathcal{F}(\mathcal{H})$, let $\xi_1, \ldots, \xi_n$ be an orthonormal basis of the (finite-dimensional) subspace $T\mathcal{H}$. Then, given $\xi \in \mathcal{H}$, there exist (unique) coefficients $a_1, \ldots, a_n$ such that $T\xi = a_1\,\xi_1 + \cdots + a_n\,\xi_n$. These coefficients are very easy to identify, for $\langle T\xi, \xi_j\rangle = a_j$. Because $T$ is bounded, the map $\xi \longmapsto \langle T\xi, \xi_j\rangle$ is bounded; by the Riesz Representation Theorem 4.5.4 there exist $\eta_1', \ldots, \eta_n' \in \mathcal{H}$ with $\langle T\xi, \xi_j\rangle = \langle \xi, \eta_j'\rangle$ for all $\xi \in \mathcal{H}$. Then $T\xi = \sum_j \langle \xi, \eta_j'\rangle\,\xi_j$. If we now apply the Gram–Schmidt orthonormalization process to $\eta_1', \ldots, \eta_n'$, we get orthonormal vectors $\eta_1, \ldots, \eta_n$, such that each $\eta_j'$ is a linear combination of $\eta_1, \ldots, \eta_n$; then we can write

$$T\xi = \sum_{k=1}^{n}\sum_{j=1}^{n} c_{kj}\,\langle \xi, \eta_k\rangle\,\xi_j, \qquad \xi \in \mathcal{H}, \tag{10.9}$$

for certain coefficients $\{c_{kj}\}$ and orthonormal families $\{\xi_j\}$ and $\{\eta_k\}$. We will see in subsection 10.6.6 below that (10.9) can be improved a bit, in that one can require a single sum, with the coefficients non-negative; that is the Singular Value Decomposition (see Theorem 10.6.26 for the compact case).

The above shows that rank-one operators are of the form $S\nu = \langle \nu, \eta\rangle\,\xi$, and that every finite-rank operator is a sum of rank-one operators. It is common, though not universal, to write the operator $S$ as $S = \xi \otimes \eta$ (physicists would write $|\xi\rangle\langle\eta|$, though we do not advocate for the bra-ket notation here). We will use the simpler notation $\xi\eta^*$, that agrees with the standard use in $\mathbb{C}^n$. We have

$$(\xi_1\eta_1^*)(\xi_2\eta_2^*) = \xi_1(\eta_1^*\xi_2)\eta_2^* = \langle \xi_2, \eta_1\rangle\,\xi_1\eta_2^* \tag{10.10}$$

and

$$(\xi\eta^*)^* = \eta\xi^*$$

(Exercise 10.6.4).

***10.6.1. Proposition.***

*Let $T \in \mathcal{B}(\mathcal{H})$. Then the following statements are equivalent:*

*(i)* $T \in \mathcal{F}(\mathcal{H})$;

*(ii) there exist* $\eta_1, \ldots, \eta_n, \xi_1, \ldots, \xi_n \in \mathcal{H}$ *such that* $T = \sum_{k=1}^{n} \xi_k\eta_k^*$;

*(iii)* *there exist* $\eta_1, \ldots, \eta_n \in \mathcal{H}$ *and an orthonormal set* $\{\xi_1, \ldots, \xi_n\} \subset \mathcal{H}$ *such that* $T = \sum_{k=1}^{n} \xi_k \eta_k^*$;

*(iv)* *there exist* $\xi_1, \ldots, \xi_n \in \mathcal{H}$ *and an orthonormal set* $\{\eta_1, \ldots, \eta_n\} \subset \mathcal{H}$ *such that* $T = \sum_{k=1}^{n} \xi_k \eta_k^*$;

*(v)* *there exist orthonormal sets* $\{\xi_1, \ldots, \xi_n\}$ *and* $\{\eta_1, \ldots, \eta_n\} \subset \mathcal{H}$, *and coefficients* $\{c_{kj}\} \subset \mathbb{C}$ *such that*

$$T = \sum_{k=1}^{n} \sum_{j=1}^{n} c_{kj} \xi_j \eta_k^*.$$

*Proof.* Exercise 10.6.5. □

In language to be developed in Section 12.1, the next result says that $\mathcal{F}(\mathcal{H})$ is sot-dense in $\mathcal{B}(\mathcal{H})$.

**10.6.2. Proposition.**

*The set* $\mathcal{F}(\mathcal{H})$ *of all finite-rank operators is a (two-sided)* $*$*-ideal of* $\mathcal{B}(\mathcal{H})$. *Every* $T \in \mathcal{B}(\mathcal{H})$ *is a pointwise limit of operators in* $\mathcal{F}(\mathcal{H})$.

*Proof.* We know that $\mathcal{F}(\mathcal{H})$ is a (possibly non-closed ideal) of $\mathcal{B}(\mathcal{H})$ by Exercise 9.6.6. The adjoint of a finite-rank operator is finite-rank by Theorem 9.6.12, though a simple ad-hoc proof works for operators on Hilbert spaces (Exercise 10.6.8).

Given $T \in \mathcal{B}(\mathcal{H})$, $\xi \in \mathcal{H}$, fix an orthonormal basis $\{\xi_j\} \subset \mathcal{H}$. Then, by Theorem 4.3.21, $T\xi = \sum_j \langle T\xi, \xi_j \rangle \xi_j$, and $\sum_j |\langle T\xi, \xi_j \rangle|^2 < \infty$. Let $P_k$ be the orthogonal projection onto the span of $\xi_1, \ldots, \xi_k$. Then $\|T\xi - P_k T\xi\| \to 0$ (Exercise 10.6.9), and $P_k T \in \mathcal{F}(\mathcal{H})$. □

A more specific form of Proposition 10.6.2 is considered in Exercise 10.6.25.

**10.6.3. Proposition.**

*Let* $T \in \mathcal{F}(\mathcal{H})$. *Then* $\dim \operatorname{ran} T^* = \dim \operatorname{ran} T$.

*Proof.* The polar decomposition of $T$ gives us a partial isometry $V : \operatorname{ran} T^* \to \operatorname{ran} T$. As it is injective on $\operatorname{ran} T^*$ it will preserve dimension, so $\dim \operatorname{ran} T^* \leq \dim \operatorname{ran} T$. But we can do the same for $T^*$ instead of $T$, so $\dim \operatorname{ran} T^* = \dim \operatorname{ran} T$. □

**10.6.2. Limits of finite-rank operators.** We recall that a linear operator $T : \mathcal{H} \to \mathcal{K}$ is **compact** if $\overline{T(B_1^{\mathcal{H}}(0))}$ is compact in $\mathcal{K}$. It follows straight from the definition that a compact operator is bounded.

It follows from Corollary 9.8.18 that when $\mathcal{H}$ is infinite-dimensional and separable (and hence isomorphic to $\ell^2(\mathbb{N})$), $\mathcal{K}(\mathcal{H})$ is the unique proper ideal of $\mathcal{B}(\mathcal{H})$. Since Hilbert spaces are very well-behaved Banach spaces, from Section 9.6 we get

***10.6.4. Proposition.***

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. The following statements are equivalent:*

*(i) $T$ is compact;*

*(ii) $T \in \overline{\mathcal{F}(\mathcal{H}, \mathcal{K})}$;*

*(iii) $T$ is completely continuous;*

*(iv) $T$ is strictly singular.*

*Proof.* (i) $\iff$ (ii) Proposition 9.8.5, as orthonormal bases are Schauder.

(i) $\iff$ (iii) This is Proposition 9.6.5, since $\mathcal{H}$ is reflexive.

(i) $\implies$ (iv) By Exercise 9.8.15 we know that $\mathcal{SS}(\mathcal{H}, \mathcal{K})$ is a closed, two-sided ideal of $\mathcal{B}(\mathcal{H}, \mathcal{K})$. It is proper when $\dim \mathcal{H} = \infty$, for the identity $I$ is not strictly singular. Then $\mathcal{SS}(\mathcal{H}, \mathcal{K}) = \mathcal{K}(\mathcal{H}, \mathcal{K})$ by Corollary 9.8.18. □

*10.6.5. Examples.*

(i) Every finite-rank operator $T$ is compact.

(ii) (A non finite-rank compact operator) Fix a countable orthonormal set $\{\xi_n\}_n \subset \mathcal{H}$ and define $T\xi_n = (1/n)\xi_n$. This can be extended by linearity and continuity to a bounded operator $T \in \mathcal{B}(\mathcal{H})$, that is a limit of finite-rank operators (Exercise 10.6.6). The operator $T$ cannot be finite rank because the whole set $\{\xi_n\}$ is contained in its image (so it's range is actually dense if $\{\xi_n\}$ is a basis).

(iii) (A non-compact operator) Let $I_{\mathcal{H}}$ be the identity operator, $I_{\mathcal{H}}\xi = \xi$ for every $\xi \in \mathcal{H}$. Then $I_{\mathcal{H}}$ is not compact, provided that $\mathcal{H}$ is infinite-dimensional, by Theorem 5.2.9. A second argument to justify that $I_{\mathcal{H}}$

is not compact is given by Lemma 9.6.3). A third argument runs as follows. Let $T \in \mathcal{B}(\mathcal{H})$ be any finite-rank operator. Necessarily, the kernel of $T$ is not trivial (Exercise 10.6.11); let $\xi \in \ker T$ with $\|\xi\| = 1$. Then $\|(I_{\mathcal{H}} - T)\xi\| = \|\xi\| = 1$, so $\|I_{\mathcal{H}} - T\| \geq 1$. We conclude that for any $T \in \mathcal{K}(\mathcal{H})$, $\|I_{\mathcal{H}} - T\| \geq 1$.

**10.6.3. The spectrum of a compact operator.** As opposed to arbitrary operators on an infinite-dimensional Hilbert space, the spectrum of a compact operator is easily understood. We already have the basic description from Theorem 9.6.13.

*10.6.6. Remark*. We know from Lemma 9.6.3 that a compact operator on an infinite-dimensional Hilbert space cannot be surjective. It can have dense range, though. Fix an orthonormal basis $\{\xi_j\} \subset \mathcal{H}$. Let $T \in \mathcal{K}(\mathcal{H})$ be the operator defined by $T\xi_j = (1/j)\xi_j$. Then the range of $T$ is not closed. Indeed, since $\xi_j \in \operatorname{ran} T$ for all $j$, the range of $T$ is dense. Now consider $\xi \in \mathcal{H}$ given by $\xi = \sum_j (1/j)\xi_j$. Then $\xi$ cannot be in the range of $T$.

*10.6.7. Remark*. By Theorem 9.6.13, the spectrum of $T \in \mathcal{K}(\mathcal{H})$ has to fall into one of the following three situations:

(i) $\sigma(T) = \{0\}$;

(ii) $\sigma(T) = \{0, \lambda_1, \ldots, \lambda_n\}$;

(iii) $\sigma(T) = \{0\} \cup \{\lambda_1, \lambda_2, \ldots\}$, where $\lim_{k\to\infty} \lambda_k = 0$.

In any of the three situations, $\ker(T - \lambda_j I)$ is finite-dimensional if $\lambda_j \neq 0$. Exercises (10.6.12), (10.6.13), and (10.6.14) show that all three situations can occur for arbitrary choices of $\lambda_j$.

We re-state Corollary 9.6.14:

***10.6.8. Theorem. (Fredholm Alternative)***

*Let $T \in \mathcal{K}(\mathcal{H})$, $\lambda \neq 0$. Then $T - \lambda I$ is either invertible, or it has nontrivial finite-dimensional kernel.*

**10.6.4. Selfadjoint and normal compact operators.** We saw in Proposition 10.1.8 that $\sigma(T^*) = \overline{\sigma(T)}$. In particular, when $T \in \mathcal{K}(\mathcal{H})$ and $\mu$ is a nonzero eigenvalue for $T$, then $\bar{\mu}$ is an eigenvalue for $T^*$. But this relation does not extend to eigenvectors. For instance let

$$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \eta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then $T\eta = \eta$, but $T^*\eta = 0$. Things are better when $T$ is normal, though:

***10.6.9. Lemma.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal, and $\mu \in \sigma(T)$ be an eigenvalue with eigenvector $\eta$. Then $\bar{\mu} \in \sigma(T^*)$ is an eigenvalue for $T^*$ with eigenvector $\eta$.*

*Proof.* If $T\eta = \mu\eta$, then $(T - \lambda I)\eta = 0$. Note that $T - \mu I$ is normal. Then $\|(T^* - \mu I)\eta\| = \|(T - \mu I)^*\eta\| = \|(T - \mu I)\eta\| = 0$ (Exercise 10.6.1). Thus $T^*\eta = \mu\eta$. □

Even when $T$ is not normal, the fact that all nonzero elements of the spectrum of a compact operator are eigenvalues guarantee that $T$ and $T^*$ have the same eigenvalues. But the same is not necessarily true if 0 is an eigenvalue; it is possible that $0 \in \sigma_p(T)$ but $0 \notin \sigma_p(T^*)$ (Exercise 10.6.6).

***10.6.10. Lemma.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\lambda, \mu \in \sigma(T)$ be distinct eigenvalues. If $\xi, \eta$ are eigenvectors for $\lambda, \mu$ respectively, then $\langle \xi, \eta \rangle = 0$.*

*Therefore if $P, Q$ are the orthogonal projections onto* $\ker(T - \lambda I)$ *and* $\ker(T - \mu I)$ *respectively, then $PQ = 0$.*

*Proof.* We have, using Lemma 10.6.9

$$\lambda\langle \xi, \eta \rangle = \langle \lambda\xi, \eta \rangle = \langle T\xi, \eta \rangle = \langle \xi, T^*\eta \rangle = \langle \xi, \bar{\mu}\eta \rangle = \mu\langle \xi, \eta \rangle.$$

As $\lambda \neq \mu$, we must have $\langle \xi, \eta \rangle = 0$.

As for the projections $P$ and $Q$, we have just proven that they have orthogonal ranges, so for any $\xi \in \mathcal{H}$ we have $\langle PQ\xi, \xi \rangle = \langle Q\xi, P\xi \rangle = 0$, and then $PQ = 0$. □

**10.6.11. Lemma.**

*Let $\lambda \in c_0$, and $\{P_k\} \subset \mathcal{B}(\mathcal{H})$ a sequence of pairwise orthogonal finite-rank projections. Then the series*

$$\sum_{k=1}^{\infty} \lambda_k P_k$$

*converges in norm, and it defines an operator $T \in \mathcal{K}(\mathcal{H})$ with $\|T\| = \|\lambda\|_\infty$ and $\sigma(T) = \{0\} \cup \{\lambda_k : k\}$.*

*Proof.* We need to show that the tails of the series are Cauchy. For $\xi \in \mathcal{H}$,

$$\begin{aligned}\Big\|\sum_{k=m}^{n} \lambda_k P_k\xi\Big\|^2 &= \sum_{k,j=m}^{n} \lambda_k\overline{\lambda_j}\,\langle P_k\xi, P_j\xi\rangle = \sum_{k=m}^{n} |\lambda_k|^2\,\|P_k\xi\|^2 \\ &\leq \max\{|\lambda_k|^2 : m \leq k \leq n\} \sum_{k=m}^{n} \|P_k\xi\|^2 \\ &\leq \max\{|\lambda_k|^2 : m \leq k \leq n\}\,\|\xi\|.\end{aligned}$$

As $\lambda \in c_0$, given $\varepsilon > 0$ there exists $m_0$ such that $|\lambda_k| < \varepsilon$ for all $k \geq m_0$. Then, for all $n \geq m \geq m_0$, the above shows that

$$\Big\|\sum_{k=m}^{n} \lambda_k P_k\Big\| \leq \varepsilon,$$

and so the series converges in norm. Then $T \in \overline{\mathcal{F}(\mathcal{H})} = \mathcal{K}(\mathcal{H})$.

Since $\sum_{k=1}^{m} P_k$ is a projection, the above estimates also give us

$$\begin{aligned}\|T\xi\|^2 = \Big\|\sum_{k=1}^{m} \lambda_k P_k\xi\Big\|^2 &\leq \|\lambda\|_\infty^2 \Big\|\sum_{k=1}^{m} P_k\xi\Big\|^2 \\ &= \|\lambda\|_\infty^2 \Big\|\Big(\sum_{k=1}^{m} P_k\Big)\xi\Big\|^2 \leq \|\lambda\|_\infty^2\,\|\xi\|.\end{aligned}$$

Then $\|T\| \leq \|\lambda\|_\infty$. If $k$ is such that $\|\lambda\|_\infty = |\lambda_k|$, then with $\xi \in P_k\mathcal{H}$ a unit vector, $\|T\xi\| = \|\lambda_k\xi\| = \|\lambda\|_\infty$. Hence $\|T\| = \|\lambda\|_\infty$.

As for the spectrum, we can repeat the ideas from Exercise 9.5.8. Namely, we have $P_k\mathcal{H} \subset \ker(T - \lambda_k I)$, so $\lambda_k \in \sigma(T)$ for all $k$. Hence $\{0\} \cup \{\lambda_k : k\} = \overline{\{\lambda_k : k\}} \subset \sigma(T)$. Conversely, if $\lambda \notin \overline{\{\lambda_k : k\}}$, there exists $\delta > 0$ with $|\lambda - \lambda_k| \geq \delta$ for all $k$. With the same argument as the first part of the proof we can construct, with $Q = I - \sum_k P_k$,

$$S = -\frac{1}{\lambda}\,Q + \sum_{k=1}^{\infty} \frac{1}{\lambda - \lambda_k}\,P_k.$$

As above, the series converges in norm and $\|S\| = \max\left\{\frac{1}{|\lambda-\lambda_k|} : k\right\}$. Computing,

$$(T - \lambda I) = S(T - \lambda I) = Q + \sum_{k=1}^{\infty} P_k = I.$$

Hence $\lambda \notin \sigma(T)$, showing that $\sigma(T) = \{0\} \cup \{\lambda_k : k\}$. □

Here is the big deal result that we can get for compact operators on a Hilbert space, due to the availability of orthogonal projections:

**10.6.12. Theorem.** ***(Spectral Theorem for Compact Operators)***

*Let $T \in \mathcal{K}(\mathcal{H})$ be normal. Then there exist sequences $\{\lambda_k\} \subset \mathbb{C}$ and $\{P_k\}_{k=1}^{\infty}$ of pairwise orthogonal finite-rank projections, such that*

$$T = \sum_{k=0}^{\infty} \lambda_k P_k, \tag{10.11}$$

*where $\lambda_0 = 0$ and $P_0$ is the orthogonal projection onto $\ker T$ (possibly zero). The sequence $\{\lambda_k\}$ converges to zero, the series converges in norm, $\sigma(T) = \{\lambda_k : k \in \mathbb{N} \cup \{0\}\}$, and $\overline{\bigcup_k P_k\mathcal{H}} = \mathcal{H}$. When $T$ is selfadjoint $\lambda_k \in \mathbb{R}$ for all $k$.*

*Proof.* Using Theorem 9.6.13, let $\{\lambda_k\}$ be an enumeration of $\sigma(T) \setminus \{0\}$; we know from the theorem that $\lambda_k \to 0$. Let $P_k$ be the orthogonal projection onto $\ker(T - \lambda_k I)$. Again by Theorem 9.6.13, each $P_k$ is finite-rank. By Lemma 10.6.10, $P_k P_j = 0$ when $k \neq j$. Then the operator $S = \sum_{k=0}^{\infty} \lambda_k P_k$ in (10.11) exists by Lemma 10.6.11.

If $\xi \in P_k\mathcal{H}$, then

$$T\xi = \lambda_k \xi = \lambda_k P_k \xi = S\xi.$$

The proof will be complete if we show that $T$ is zero on the orthogonal complement of $\operatorname{span} \bigcup_{k\geq 1} P_k\mathcal{H}$. Let

$$\mathcal{K} = \left[\bigcup_k P_k\mathcal{H}\right]^{\perp}.$$

If $\xi \in \mathcal{K}$ and $\eta \in P_j\mathcal{H}$, then $T\eta = \lambda_j\eta$ and $T^*\eta = \overline{\lambda_j}\eta$ by Lemma 10.6.9. Then

$$\langle \eta, T\xi\rangle = \langle T^*\eta, \xi\rangle = \overline{\lambda_j}\langle \eta, \xi\rangle = 0.$$

Hence $T\xi \in \mathcal{K}$, showing that $\mathcal{K}$ is an invariant subspace for $T$; a similar computation shows that $\mathcal{K}$ is also invariant for $T^*$. Then $T|_{\mathcal{K}}$ is normal. If

$T|_{\mathcal{K}} \neq 0$, by Proposition 10.3.3 there is nonzero $\lambda \in \sigma(T)$. And by Theorem 9.6.13, there exists $k \in \mathbb{N}$ with $\lambda \in P_k\mathcal{H}$; but this is a contradiction because an eigenvector for $\lambda$ would lie in $\mathcal{K} \cap \mathcal{K}^\perp = \{0\}$. So $T|_{\mathcal{K}} = 0$. □

It is instructive to write (10.11) in the case where $T$ is a diagonal matrix.

*10.6.13. Remark*. In Theorem 10.6.12 we expressed the normal operator $T$ in such a way that all $\lambda_k$ are pairwise distinct. If we write each $P_k = P_{k,1} + \cdots + P_{k,m(k)}$ as a sum of pairwise orthogonal rank-one projections, we get to write

$$T = \sum_k \sum_{j=1}^{m(k)} \lambda_k \, P_{k,j} = \sum_k \lambda_k' \, P_k',$$

where now each $P_k'$ is a rank-one projection, and there might be repetitions among the $\lambda_k'$. It will sometimes be more useful to express $T$ this way. We say that $\{\lambda_k'\}$ are the eigenvalues of $T$ **counting multiplicities**.

The Spectral Theorem simplifies dealings with normal compact operators in many ways, as we will see in the next results. For starters, compare the proof of Corollary 10.6.14 with that of Proposition 10.4.4 (granted the former only works for compact positive operators, while the latter for all positive operators; but the point still stands).

***10.6.14. Corollary.***

*Let $T \in \mathcal{K}(\mathcal{H})$ be positive. If we express $T$ as in (10.11), then*

$$T^{1/2} = \sum_{k=1}^\infty \lambda_k^{1/2} \, P_k.$$

*Proof.* Since the projections $P_k$ are pairwise orthogonal, a straightforward computation shows that $(T^{1/2})^2 = T$. As the positive square root is unique (this was shown in Proposition 10.4.4), this is enough. □

***10.6.15. Corollary.***

*Let $T \in \mathcal{K}(\mathcal{H})$. Then $\overline{\operatorname{ran} T}$ is separable.*

*Proof.* Assume first that $T \geq 0$. Then we can write $T = \sum_{k=0}^{\infty} \lambda_k P_k$ by the Spectral Theorem. It follows that for any $\xi \in \mathcal{H}$ we have

$$T\xi \in \overline{\operatorname{span}}^{\|\cdot\|} \bigcup_{k=1}^{\infty} P_k\mathcal{H}.$$

Since each $P_k\mathcal{H}$ is finite-dimensional, this shows that $\operatorname{ran} T$ is spanned by countably many vectors. For arbitrary $T$, write $T = V|T|$ via the polar decomposition, and now $\operatorname{ran} T \subset V \operatorname{ran}|T|$, again spanned by a countable set. □

**10.6.16. Corollary.**

*Let $T \in \mathcal{K}(\mathcal{H})$ selfadjoint. Then there exist $T^+, T^- \in \mathcal{K}(\mathcal{H})$, both positive, such that $T = T^+ - T^-$ and $T^+T^- = 0$.*

*In particular, $\mathcal{K}(\mathcal{H}) = \operatorname{span}\{T \in \mathcal{K}(\mathcal{H}) : T \geq 0\}$.*

*Proof.* By the Spectral Theorem (10.6.12) we can write $T = \sum_{k=0}^{\infty} \lambda_k P_k$, where each $\lambda_k$ is an eigenvalue and $\lambda \in c_0$. Because $T = T^*$ we have $\lambda_k \in \mathbb{R}$ for all $k$ (Proposition 10.3.3). Let $K^+ = \{k \in \mathbb{N} : \lambda_k \geq 0\}$, $K^- = \mathbb{N} \setminus K^+$. Then we can define

$$T^+ = \sum_{k\in K^+} \lambda_k P_k, \qquad T^- = -\sum_{k\in K^-} \lambda_k P_k.$$

Both $T^+, T^-$ exist and are positive by Lemma 10.6.11 and they are compact by Proposition 10.6.4. We have $T = T^+ - T^-$ by construction. Finally, as any $T$ is a linear combination of its real and imaginary parts which are selfadjoint, it follows that $T$ is a linear combination of four positive compact operators. □

Since compact operators always have 0 in their spectrum, we know from Proposition 9.2.15 that $\sigma(T^*T) = \sigma(TT^*)$ for any $T \in \mathcal{K}(\mathcal{H})$. But we can say a bit more:

**10.6.17. Corollary.**

*Let $T \in \mathcal{K}(\mathcal{H})$. Then $\lambda \in \sigma(T^*T)$ if and only if $\lambda \in \sigma(TT^*)$, and in both cases the multiplicity is the same. In other words, if we write $T^*T$ and $T^*T$ as in (10.11),*

$$T^*T = \sum_{k=1}^{\infty} \lambda_k P_k, \qquad TT^* = \sum_{k=1}^{\infty} \lambda_k Q_k,$$

*where* $\{\lambda_k\} = \sigma(T^*T)\setminus\{0\} = \sigma(TT^*)\setminus\{0\}$ *and* $\dim Q_k\mathcal{H} = \dim P_k\mathcal{H}$ *for all* $k \in \mathbb{N}$.

*Proof.* Fix $k$. Let $\eta_1,\ldots,\eta_m$ be a basis of $P_k\mathcal{H}$. Then $T^*T\eta_j = \lambda_k\eta_j$ for all $j$. As $\ker T = \ker T^*T$ we have $T\eta_j \neq 0$, and

$$TT^*(T\eta_j) = T(T^*T\eta_j) = \lambda_k T\eta_j,$$

so $\lambda_k \in \sigma(TT^*)$. If $\sum_j a_jT\eta_j = 0$, then

$$0 = \sum_j a_jT^*T\eta_j = \sum_j a_j\lambda_k\,\eta_j.$$

As the $\eta_j$ are linearly independent, we get that $a_j\lambda_k = 0$ for all $j$, and hence $a_j = 0$ since $\lambda_k \neq 0$. Thus $T\eta_1,\ldots,T\eta_m$ are linearly independent. This shows that $\dim Q_k\mathcal{H} \geq \dim P_k\mathcal{H}$. The roles of $T$ and $T^*$ can be reversed, so $\dim Q_k\mathcal{H} = \dim P_k\mathcal{H}$. $\square$

It was no mistake that we left 0 out of Corollary 10.6.17, for the equality $\dim Q_0\mathcal{H} = \dim P_0\mathcal{H}$ can fail. Indeed, fix an orthonormal basis $\{\xi_n\}$ and consider the compact operator given by $T\xi = \sum_k \frac{1}{k}\langle \xi,\xi_n\rangle\,\xi_{n+1}$. Then $\ker T^*T = \{0\}$, while $\ker TT^* = \mathbb{C}\,\xi_1$ (Exercise 10.6.7).

There is no difficulty in checking that, for any $f \in \mathbb{C}[x]$ with $f(0) = 0$ and $T$ normal,

$$f(T) = \sum_{k=0}^{\infty} f(\lambda_k)\,P_k. \tag{10.12}$$

This suggests that we could use (10.12) to define $f(T)$ for any normal operator $T$ and $f : \sigma(T) \to \mathbb{C}$, with $f(0) = 0$ and continuous at 0 (for we need $f(\lambda_k) \to 0$ to get a compact operator). There is no issue with this, and it agrees with the continuous and Borel functional calculus that we will develop in Section 11.2 and Section 12.4. And we can prove with the tools at our disposal is that this definition is coherent:

***10.6.18. Corollary.***

*Let* $T \in \mathcal{K}(\mathcal{H})$ *be normal, and* $f : \mathbb{C} \to \mathbb{C}$, *with* $f(0) = 0$ *and continuous at* 0. *Let* $\{p_m\}$ *be a sequence of polynomials with* $p_m \to f$ *uniformly on* $\sigma(T)$. *Then* $\|f(T) - p_m(T)\| \to 0$.

*Proof.* Via the Spectral Theorem write

$$T = \sum_{k=0}^{\infty} \lambda_k\,P_k.$$

Fix $\varepsilon > 0$. By hypothesis where exists $m_0$ such that $|f(t) - p_m(t)| < \varepsilon$ for all $m \geq m_0$ and all $t \in \sigma(T)$. Then, by Lemma 10.6.11,

$$\|f(T) - p_m(T)\| = \left\| \sum_{k=1}^{\infty} \left[f(\lambda_k) - p_m(\lambda_k)\right] P_k \right\|$$
$$= \sup\{|f(\lambda_k) - p_m(\lambda_k)| : \ k\} < \varepsilon. \qquad \square$$

We also have

***10.6.19. Proposition. (Spectral Mapping, Compact Version)***

*Let $T \in \mathcal{K}(\mathcal{H})$ be normal, $f : \sigma(T) \to \mathbb{C}$ continuous at $0$ with $f(0) = 0$. Then $\sigma(f(T)) = f(\sigma(T))$.*

*Proof.* Exercise 10.6.18. $\square$

The continuity requirement of $f$ in Proposition 10.6.19 can be relaxed to $f$ bounded and continuous at 0 (note that $\sigma(f(T))$ is compact, so we need $f(\sigma(T))$ to be compact). More general functional calculus notions will be discussed in Sections 11.2 and 12.4.

**10.6.5. Eigenvalue Inequalities.** The Spectral Theorem allows one to obtain useful estimates for eigenvalues. We will show a few, and we are not even scratching the surface.

**10.6.20. Definition.**

*Let $T \in \mathcal{K}(\mathcal{H})$. The eigenvalues $\{\sigma_k\}$ of $|T|$ are called the* ***singular values*** *of $T$. When more than one operator is involved, the notation $\sigma_k(T)$ will be used.*

One of many many consequences of the Spectral Theorem is that when $T$ is normal, $\sigma_k(T) = |\lambda_k(T)|$ (Exercise 10.6.17). And, whether $T$ is normal or not, we always have the following very simple but useful equality:

***10.6.21. Lemma.***

*Let $T \in \mathcal{K}(\mathcal{H})$, $k \in \mathbb{N}$. Then $\sigma_k(T) = \sigma_k(T^*)$.*

*Proof.* We have, via Propositions 9.2.9 and 9.2.15,

$$\sigma_k(T) = \lambda_k(|T|) = \lambda_k(|T|^2)^{1/2} = \lambda_k(T^*T)^{1/2} = \lambda_k(TT^*)^{1/2}$$

$$= \lambda_k(|T^*|^2)^{1/2} = \lambda_k(|T^*|) = \sigma_k(T^*). \qquad \square$$

***10.6.22. Lemma.***

*Let $\mathcal{H}$ be an infinite-dimensional Hilbert space, $\mathcal{K}, \mathcal{N} \subset \mathcal{H}$ closed subspaces with $\dim \mathcal{N}^\perp = k$ and $\mathcal{K} \cap \mathcal{N} = \{0\}$. Then $\dim \mathcal{K} \le k$.*

*Proof.* Since $\dim \mathcal{N}^\perp = k$, we have $\dim \mathcal{H}/\mathcal{N} = k$, since $\mathcal{H}/\mathcal{N} \simeq \mathcal{N}^\perp$ (Exercise 5.3.9). Suppose that $\eta_1, \dots, \eta_r \in \mathcal{K}$ are linearly independent. If $c_1\eta_1 + \cdots + c_r\eta_r + \mathcal{N} = \mathcal{N}$ in $\mathcal{H}/\mathcal{N}$, then $c_1\eta_1 + \cdots + c_r\eta_r \in \mathcal{N}$ and this means that $c_1\eta_1 + \cdots + c_r\eta_r = 0$ and then $c_1 = \cdots = c_r = 0$ by the linear independence in $\mathcal{K}$. As $\mathcal{H}/\mathcal{N}$ admits at most $k$ linearly independent elements, this means that $\dim \mathcal{K} \le k$. $\square$

***10.6.23. Corollary. (The Min-Max Principle)***

*Let $T \in \mathcal{K}(\mathcal{H})$ be positive, and $S \in \mathcal{K}(\mathcal{H})$. Then*

$$\lambda_{k+1}(T) = \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2} = \min_{\dim \mathcal{K}=k} \max_{\xi \in \mathcal{K}^\perp} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2}, \tag{10.13}$$

*and*

$$\sigma_{k+1}(S) = \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \frac{\|S\xi\|}{\|\xi\|} = \min_{\dim \mathcal{K}=k} \max_{\xi \in \mathcal{K}^\perp} \frac{\|S\xi\|}{\|\xi\|}. \tag{10.14}$$

*Proof.* Via the Spectral Theorem we have

$$T = \sum_{k=1}^\infty \lambda_k(T)\, P_k,$$

where the projections $P_k$ are rank-one, pairwise orthogonal, and such that we have an orthonormal basis $\{\xi_k\}$ with $P_k\xi_k = \xi_k$ for all $k$. Let $\mathcal{K} \subset \mathcal{H}$ be a subspace with $\dim \mathcal{K} = k+1$. Let $\mathcal{N} = \overline{\operatorname{span}}^{\|\cdot\|}\{\xi_{k+1}, \xi_{k+2}, \dots\}$, so $\dim \mathcal{N}^\perp = k$. By Lemma 10.6.22 this means that $\mathcal{K} \cap \mathcal{N} \ne \{0\}$. Let $\xi \in \mathcal{K} \cap \mathcal{N}$ with $\|\xi\| = 1$. Then $\xi = \sum_{j=k+1}^\infty c_j\xi_j$ and

$$\langle T\xi, \xi\rangle = \sum_{j=k+1}^\infty \lambda_j(T)\,|c_j|^2 \le \lambda_k(T) \sum_{j=k+1}^\infty |c_j|^2 = \lambda_{k+1}(T).$$

Thus

$$\min_{\xi \in \mathcal{K}} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2} \le \lambda_{k+1}(T).$$

Choosing $\mathcal{K} = \operatorname{span}\{\xi_1, \ldots, \xi_{k+1}\}$ we see that the bound $\lambda_{k+1}(T)$ can be attained, and so

$$\lambda_{k+1}(T) = \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2}.$$

Now fix $\mathcal{K} \subset \mathcal{H}$ a subspace with $\dim \mathcal{K} = k$. Let $\mathcal{N} = \overline{\operatorname{span}}^{\|\cdot\|}\{\xi_1, \ldots, \xi_k\}$. Another application of Lemma 10.6.22 implies that $\mathcal{K}^\perp \cap \mathcal{N} \neq \{0\}$. If $\xi \in \mathcal{K}^\perp \cap \mathcal{N}$ with $\|\xi\| = 1$, we may write $\xi = \sum_{j=1}^k c_j \xi_j$ with $\sum_j |c_j|^2 = 1$. Then

$$\langle T\xi, \xi\rangle = \sum_{j=1}^k \lambda_j(T)\, |c_j|^2 \geq \lambda_k(T) \sum_{j=1}^k |c_j|^2 = \lambda_k(T).$$

Thus

$$\max_{\xi \in \mathcal{K}^\perp} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2} \geq \lambda_k(T).$$

Choosing $\mathcal{K} = \operatorname{span}\{\xi_1, \ldots, \xi_k\}$ we see that the bound $\lambda_{k+1}(T)$ can be attained, and so

$$\lambda_k(T) = \min_{\dim \mathcal{K}=k} \max_{\xi \in \mathcal{K}^\perp} \frac{\langle T\xi, \xi\rangle}{\|\xi\|^2}.$$

To show (10.14), from Corollary 10.6.23,

$$\begin{aligned}
\sigma_k(S) = \lambda_k(|S|) = \lambda_k(|S|^2)^{1/2} &= \max_{\dim \mathcal{K}=k+1} \left( \min_{\xi \in \mathcal{K}} \frac{\langle |S|^2\xi, \xi\rangle}{\|\xi\|^2} \right)^{1/2} \\
&= \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \left( \frac{\|\,|S|\xi\|^2}{\|\xi\|^2} \right)^{1/2} = \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \frac{\|\,|S|\xi\|}{\|\xi\|} \\
&= \max_{\dim \mathcal{K}=k+1} \min_{\xi \in \mathcal{K}} \frac{\|\,S\xi\|}{\|\xi\|}
\end{aligned}$$

since $\|S\xi\| = \|\,|S|\xi\|$ holds always. The second equality follows from (10.13) with the same idea. $\square$

***10.6.24. Corollary.***

*Let $T \in \mathcal{K}(\mathcal{H})$ and $S \in \mathcal{B}(\mathcal{H})$, $k \in \mathbb{N}$. Then*

$$\max\left\{\sigma_k(ST), \sigma_k(TS)\right\} \leq \|S\|\,\sigma_k(T).$$

*Proof.* Using (10.14),

$$\sigma_k(ST) = \min_{\dim \mathcal{K}=k} \max_{\xi \in \mathcal{K}^\perp} \frac{\|\,ST\xi\|}{\|\xi\|} \leq \|S\| \min_{\dim \mathcal{K}=k} \max_{\xi \in \mathcal{K}^\perp} \frac{\|\,T\xi\|}{\|\xi\|} = \|S\|\,\sigma_k(T).$$

And now by Lemma 10.6.21 and the above inequality

$$\sigma_k(TS) = \sigma_k(S^*T^*) \leq \|S^*\|\,\sigma_k(T^*) = \|S\|\,\sigma_k(T). \qquad \square$$

It is important to mention that while $ST$ and $TS$ have the same nonzero eigenvalues by Proposition 9.2.15, the same is not true for singular values. This already shows up in dimension 2, where we can take $S = E_{11}$, $T = E_{12}$ in $M_2(\mathbb{C})$, and then $\sigma_1(TS) = \sigma_2(TS) = 0$ since $TS = 0$, while $\sigma_1(ST) = \sigma_1(T) = 1$.

***10.6.25. Corollary.***

*Let $S, T \in \mathcal{K}(\mathcal{H})$ and $k, j \in \mathbb{N} \cup \{0\}$. Then*

$$\sigma_{k+j+1}(S+T) \le \sigma_{k+1}(S) + \sigma_{j+1}(T).$$

*Proof.* From Corollary 10.6.23

$$\begin{aligned}
\sigma_{k+j+1}(S+T) &= \min_{\dim\mathcal{K}=k+j}\ \max_{\xi\in\mathcal{K}^\perp} \frac{\|(S+T)\xi\|}{\|\xi\|} \\
&= \min_{\dim\mathcal{K}_1=k,\ \dim\mathcal{K}_2=j} \left( \max_{\xi\in(\mathcal{K}_1+\mathcal{K}_2)^\perp} \frac{\|(S+T)\xi\|}{\|\xi\|} \right) \\
&\le \min_{\dim\mathcal{K}_1=k,\ \dim\mathcal{K}_2=j} \left( \max_{\xi\in(\mathcal{K}_1+\mathcal{K}_2)^\perp} \frac{\|S\xi\|}{\|\xi\|} + \max_{\xi\in(\mathcal{K}_1+\mathcal{K}_2)^\perp} \frac{\|T\xi\|}{\|\xi\|} \right) \\
&\le \min_{\dim\mathcal{K}_1=k,\ \dim\mathcal{K}_2=j} \left( \max_{\xi\in(\mathcal{K}_1)^\perp} \frac{\|S\xi\|}{\|\xi\|} + \max_{\xi\in(\mathcal{K}_2)^\perp} \frac{\|T\xi\|}{\|\xi\|} \right) \\
&= \sigma_{k+1}(S) + \sigma_{j+1}(T).
\end{aligned}$$

□

**10.6.6. The Singular Value Decomposition.** A fairly direct application of the polar decomposition gives the well-known **Singular Value Decomposition**:

***10.6.26. Theorem. (Singular Value Decomposition for Compact Operators)***

*Let $T \in \mathcal{K}(\mathcal{H})$. Then there exist non-negative numbers $\{\sigma_k\}$, with $\sigma_k \to 0$, and orthonormal sequences $\{\xi_k\}$ and $\{\eta_k\}$ in $\mathcal{H}$ such that*

$$T = \sum_{k=1}^{\infty} \sigma_k\, \eta_k \xi_k^*. \tag{10.15}$$

*That is to say,*

$$T\xi = \sum_{k=1}^{\infty} \sigma_k\, \langle \xi, \xi_k \rangle\, \eta_k, \qquad \xi \in \mathcal{H}.$$

*Proof.* Via the polar decomposition we can write $T = V|T|$, with $V : \overline{\mathrm{ran}} T^* \to \overline{\mathrm{ran}} T$ a partial isometry. By the Spectral Theorem (Theorem 10.6.12) we can write

$$|T| = \sum_{k=1}^{\infty} \sigma_k \, P_k,$$

where $\sigma_k \geq 0$ for all $k$ (since $|T| \geq 0$), $P_k$ is a rank-one projection for all $k$, and $\sigma_k$ decreases to zero. Then

$$T = \sum_{k=1}^{\infty} \sigma_k \, V P_k.$$

The operators $VP_k$ are rank-one. Choose unit vectors $\{\xi_k\}$ with $\xi_k \in P_k \mathcal{H}$. When $\sigma_k \neq 0$ we have $\xi_k = \sigma_k^{-1} |T| \xi_k \in \mathrm{ran}\, |T| \subset \overline{\mathrm{ran}} T^*$. As $V^*V$ is the orthogonal projection onto $\overline{\mathrm{ran}} T^*$, it follows that $V^*V\xi_k = \xi_k$ for all $k$ such that $\sigma_k \neq 0$. Then $\|V\xi_k\|^2 = \langle V^*V\xi_k, \xi_k\rangle = \langle \xi_k, \xi_k\rangle = \|\xi_k\|^2 = 1$, and if $k \neq j$ and $\sigma_k \sigma_j > 0$ then

$$\langle V\xi_k, V\xi_j\rangle = \langle V^*V\xi_k, \xi_j\rangle = \langle \xi_k, \xi_j\rangle = 0.$$

So the sequence $\{\eta_k\}$ given by $\eta_k = V\xi_k$, is orthonormal. And we have

$$VP_k\xi = V(\langle \xi, \xi_k\rangle \, \xi_k = \langle \xi, \xi_k\rangle \, \eta_k.$$

Hence

$$T\xi = \sum_{k=1}^{\infty} \sigma_k \, V P_k = \sum_{k=1}^{\infty} \sigma_k \, \langle \xi, \xi_k\rangle \, \eta_k.$$

That is,

$$T = \sum_{k=1}^{\infty} \sigma_k \, \eta_k \xi_k^*.$$ □

***10.6.27. Corollary.***

*Let $T \in \mathcal{K}(\mathcal{H})$ and $\{\sigma_k\}$ its singular values. Then $\|T\| = \sigma_1$.*

*Proof.* If we write the singular value decomposition

$$T = \sum_{k=1}^{\infty} \sigma_k \, \eta_k \xi_k^*,$$

then $\|T\xi_1\| = \sigma_1 \, \|\eta_1\| = \sigma_1$. So $\|T\| \geq \sigma_1$. Conversely, Given $\xi \in h$ with $\|\xi\| = 1$, let $\{\nu_j\}$ be an orthonormal basis of $\{\xi_k\}^{\perp}$. Then $\{\xi_k\} \cup \{\nu_j\}$ is an orthonormal basis for $\mathcal{H}$ and

$$\xi = \sum_k c_k \xi_k + \sum_j d_j \nu_j,$$

with $\sum_k |c_k|^2 + \sum_j |d_j|^2 = 1$. Then

$$\|T\xi\|^2 = \left\| \sum_k \sigma_k c_k \xi_k \right\|^2 = \sum_k |c_k|^2 \, |\sigma_k|^2 \le \sigma_1^2 \sum_k |c_k|^2 \le \sigma_1^2.$$

Hence $\|T\| \le \sigma_1$, proving the equality. □

Corollary 10.6.27 can be proven directly from the Spectral Theorem by using that $\|T\| = \|T^*T\|^{1/2}$.

**10.6.28. Corollary.**

*The singular values are unique. Concretely, if $T$ is as in (10.15) and also $T = \sum_{k=1}^{\infty} \alpha_k \, \nu_k \rho_k^*$ with $\alpha_k \ge 0$, in non-increasing order, and $\{\nu_k\}$, $\{\rho_k\}$ orthonormal, then $\alpha_k = \sigma_k$ for each $k$.*

*Proof.* We have

$$\sum_k \sigma_k^2 \, \eta_k \eta_k^* = T^*T = \sum_k \alpha_k^2 \, \rho_k \rho_k^*.$$

By Corollary 10.6.27, $\alpha_1^2 = \|T^*T\| = \sigma_1^2$, so $\alpha_1 = \sigma_1$. We need to account for repetitions, so let $m, n$ such that $\sigma_1 = \ldots = \sigma_m > \sigma_{m+1}$ and $\alpha_1 = \cdots = \alpha_n > \alpha_{n+1}$. Choose $f : \mathbb{C} \to \mathbb{C}$ such that $f(\sigma_1) = 1$, $f(\sigma_{m+k}) = f(\alpha_{n+k}) = 0$ for all $k \ge 1$. Then

$$\sum_{k=1}^{m} \eta_k \eta_k^* = f(T^*T) = \sum_{j=1}^{n} \rho_k \rho_k^*.$$

As the operator on the left has rank $m$ and the operator on the right has rank $n$, $m = n$. Then $\sum_{k=1}^m \sigma_k \eta_k \eta_k^* = \sum_{j=1}^m \alpha_j \rho_k \rho_k^*$, which implies

$$\sum_{k=m+1}^{\infty} \sigma_k^2 \, \eta_k \eta_k^* = \sum_{k=m+1}^{\infty} \alpha_k^2 \, \rho_k \rho_k^*.$$

Now the argument can be rehashed to obtain inductively that $\alpha_k = \sigma_k$ for all $k$. □

Let us denote by $\mathcal{F}_n(\mathcal{H})$ the space of finite-rank operators on $\mathcal{H}$ of rank $n$, and $\mathcal{F}_0(\mathcal{H}) = \{0\}$.

**10.6.29. Corollary.**

*Let $T \in \mathcal{K}(\mathcal{H})$. Then, for all $k \in \mathbb{N}$,*

$$\sigma_k(T) = \min\{\|T - S\| : \ S \in \mathcal{F}_{k-1}(\mathcal{H})\}. \tag{10.16}$$

*Proof.* Let $T = \sum_{k=1}^{\infty} \sigma_k \, \eta_k \xi_k^*$ be the singular value decomposition of $T$. Fix $k$. Then $S = \sum_{j=1}^{k-1} \sigma_k \, \eta_k \xi_k^* \in \mathcal{F}_{k-1}(\mathcal{H})$ and by Corollary 10.6.27

$$\|T - S\| = \left\| \sum_{j=k}^{\infty} \sigma_k \, \eta_k \xi_k^* \right\| = \sigma_k.$$

So $\inf\{\|T - S\| : \ S \in \mathcal{F}_{k-1}(\mathcal{H})\} \leq \sigma_k$. Conversely, let $S \in \mathcal{F}_{k-1}(\mathcal{H})$. As $\dim S\mathcal{H} = k-1$, using Proposition 10.6.3 we get $\dim(\ker S)^{\perp} = \dim \operatorname{ran} S^* = \dim \operatorname{ran} S = k-1$. Then by Corollary 10.6.23

$$\sigma_k(T) \leq \max\big\{\|T\xi\| : \ \xi \in \ker S, \ \|\xi\| = 1\big\}.$$

For any $\xi \in \ker S$,

$$\|T\xi\| = \|T\xi - S\xi\| \leq \|T - S\| \, \|\xi\|.$$

Thus $\sigma_k(T) \leq \|T - S\|$, finishing the proof of (10.16). ☐

A remarkable consequence of (10.16) is that it is phrased without use of the inner product, making the definition valid for operators acting on a Banach space.

**10.6.7. The Volterra Operator.** On Example 9.6.15 we introduced the Volterra operator $V \in \mathcal{B}(L^2[0,1])$ and proved that it is compact with $\sigma(V) = \{0\}$. This shows that $V$ is not normal, for $\operatorname{spr}(V) = 0 < \|V\|$ (alternatively, it is not hard to calculate $V^*V$ and $VV^*$ explicitly, as we see below).

Here is a way to calculate $\|V\|$. First of all, it is not hard to check that the adjoint of $V$ is given by

$$V^* f(s) = \int_s^1 f(t) \, dt, \qquad f \in L^2[0,1]$$

(Exercise 10.6.20). We have $\|V\| = \|V^*V\|^2$. As $V^*V$ is compact and positive, its norm coincides with its largest eigenvalue (Corollary 10.6.27). We have

$$V^*V f(x) = \int_x^1 \int_0^t f(s) \, ds \, dt. \tag{10.17}$$

If we differentiate the equality $V^*V f(x) = \lambda f(x)$ twice (recall from Example 9.6.15 that if $\lambda \neq 0$ then $f$ is infinitely differentiable), we get the differential equation $-f(x) = \lambda f''(x)$. From (10.17) the initial conditions are $f(1) = 0$, $f'(0) = 0$. Since we know that $\lambda > 0$ this gives, up to a multiple,

$$f(x) = \cos\frac{x}{\sqrt{\lambda}}, \qquad \text{subject to } f(1) = 0.$$

Thus

$$\frac{1}{\sqrt{\lambda}} = \frac{(2k+1)\pi}{2}, \qquad k \in \mathbb{N},$$

so

$$\sqrt{\lambda} = \frac{2}{(2k+1)\pi}.$$

Using $k = 0$ to get the largest $\lambda$,

$$\|V\| = \frac{2}{\pi}.$$

**10.6.8. Exercises.**

(10.6.1) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\xi \in \mathcal{H}$. Show that $\|T^*\xi\| = \|T\xi\|$. Use an example to show that the equality is not necessarily true when $T$ is not normal.

(10.6.2) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Prove the inequality $\|T\xi\|^2 \leq \|T^2\xi\|\,\|\xi\|$, and use it to show that if $T^2$ is compact, then $T$ is compact. Show, with an example, that normality is crucial as a hypothesis for both assertions.

(10.6.3) Prove the identity (10.10).

(10.6.4) Given a rank-one operator $\xi\eta^*$, show that

$$(\xi\eta^*)^* = \eta\xi^*. \tag{10.18}$$

(10.6.5) Write a complete proof of Proposition 10.6.1.

(10.6.6) Show that the operator $T$ from (ii) in Examples 10.6.5 is well-defined, and that it is a limit of finite-rank operators.

(10.6.7) Fix an orthonormal basis $\{\xi_n\}$ for $\mathcal{H}$ and define

$$T\xi = \sum_{k=1}^{\infty} \frac{1}{k}\,\langle \xi, \xi_k\rangle\,\xi_{k+1}.$$

(i) Show that $T \in \mathcal{K}(\mathcal{H})$;

(ii) find $T^*$;

(iii) show that $\ker T^*T = \{0\}$ and that $\ker TT^* = \mathbb{C}\,\xi_1$.

(10.6.8) If $T \in \mathcal{F}(\mathcal{H})$, show that $T^* \in \mathcal{F}(\mathcal{H})$.

(10.6.9) Let $\mathcal{H}$ be a Hilbert space, $\{\eta_j\}_{j\in J}$ an orthonormal basis, and $T \in \mathcal{B}(\mathcal{H})$. For each finite $F \subset J$, let $P_F$ be the orthogonal projection onto $\operatorname{span}\{\xi_j : j \in F\}$. Show that for each $\xi \in \mathcal{H}$, $\|(T - P_F T)\xi\| \to 0$, $\|(T - TP_F)\xi\| \to 0$, and $\|(T - P_F T P_F)\xi\| \to 0$.

(10.6.10) Show that for $T \in \mathcal{K}(\mathcal{H})$ and $\{\xi_j\}$ an orthonormal basis,

$$\lim_j T\xi_j = 0.$$

(10.6.11) Prove that whenever $\mathcal{H}$ is infinite-dimensional, any finite-rank operator $T \in \mathcal{F}(\mathcal{H})$ has non-trivial kernel.

(10.6.12) Let $\mathcal{H} = \ell^2(\mathbb{N})$. Define operators

$$T(a_1, a_2, \ldots) = \big(a_1, \tfrac{a_2}{2}, \tfrac{a_3}{3}, \ldots\big),$$

$$S(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

Prove that $T$ is an injective compact operator with

$$\sigma(T) = \{0, 1, 1/2, 1/3, \ldots\},$$

and that $R = ST$ is an injective compact operator with $\sigma(R) = \{0\}$. Conclude that $Z = R^*$ is a non-injective compact operator with dense range and $\sigma(Z) = \{0\}$.

(10.6.13) Let $\mathcal{H}$ be a Hilbert space with $\dim \mathcal{H} > k$ and

$$A = \{0, \lambda_1, \ldots, \lambda_k\} \subset \mathbb{C}.$$

Show that there exists $T \in \mathcal{K}(\mathcal{H})$ with $\sigma(T) = A$.

(10.6.14) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space and $\{\lambda_k\} \subset \mathbb{C}$ be a sequence with $\lim \lambda_k = 0$. Let $A = \{0\} \cup \{\lambda_k\}$. Show that there exists $T \in \mathcal{K}(\mathcal{H})$ with $\sigma(T) = A$.

(10.6.15) What is the relation between the operator in Exercise 10.6.6 and the operator in Exercise 10.6.12?

(10.6.16) Write the explicit form of (10.11) in the case where $T$ is a diagonal matrix.

(10.6.17) Show that if $T$ is normal then $\sigma_k(T) = |\lambda_k(T)|$.

(10.6.18) Prove Proposition 10.6.19.

(10.6.19) Let $T \in \mathcal{B}(\mathcal{H})$ be a positive compact operator. Show that

$$\xi \in \operatorname{ran} T \iff \text{there exists } C > 0 \text{ such that}$$

$$|\langle \xi, \eta \rangle| \leq C \|T\eta\|, \qquad \eta \in \mathcal{H}.$$

Show by example that the condition can fail for $\xi \in \overline{\operatorname{ran} T}$.

(10.6.20) Show that if $V$ is the Volterra operator on $L^2[0,1]$, its adjoint $V^*$ is given by

$$V^* f(s) = \int_s^1 f(t)\, dt, \qquad f \in L^2[0,1].$$

(10.6.21) Let $T \in \mathcal{K}(\mathcal{H})$ and let $T = V|T|$ be its Polar Decomposition. Show that $|T| \in \mathcal{K}(\mathcal{H})$.

(10.6.22) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Use the Polar Decomposition to show that the unit ball of $\mathcal{K}(\mathcal{H})$ has no extreme points.

(10.6.23) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Show that $\mathcal{K}(\mathcal{H})$ is not a dual by using Exercise 10.6.22 and Krein–Milman (Theorem 7.5.11).

(10.6.24) Let $T \in \mathcal{K}(\mathcal{H})$ be normal. Show that $\overline{\operatorname{conv}}\, \sigma(T) = \overline{W(T)}$.

(10.6.25) Let $\mathcal{H}$ be a Hilbert space with an orthonormal basis $\{\eta_j\}_{j\in J}$. For each $k, j \in J$, let $E_{kj}$ be the rank-one operator $E_{kj}\xi = \langle \xi, \eta_j \rangle \eta_k$. These operators are called **matrix units** and satisfy the relations

$$E_{kj}E_{ab} = \delta_{j,a} E_{kb}. \qquad E_{kj}^* = E_{jk}. \tag{10.19}$$

Using notation we have also discussed, $E_{kj} = \eta_k \eta_j^*$.

(i) Prove the matrix unit relations (10.19).

(ii) Show every $E_{kj}$ is a partial isometry and $\{E_{kk}\}$ are pairwise orthogonal mutually equivalent projections.

(iii) Let $T \in \mathcal{B}(\mathcal{H})$. Show that there exist unique numbers $\{t_{kj}\}_{k,j\in J}$ such that

$$T = \sum_{k,j} t_{kj} E_{kj}, \tag{10.20}$$

where the series converges pointwise (if coming from a later chapter, the series converges sot).

## 10.7. Trace-Class Operators

Some compact operators are better than others. Or rather, certain classes of compact operators arise naturally. A central player in this setting will be the trace. In finite-dimension, this is the well-known linear functional obtained by adding the diagonal entries of a matrix representation of the operator (or equivalently, adding all the eigenvalues, counting multiplicities). This idea admits a nice generalization to infinite dimension.

### 10.7.1. Definition.

*Let $\mathcal{H}$ be a Hilbert space and fix an orthonormal basis $\{\xi_n\}$. An operator $T \in \mathcal{B}(\mathcal{H})$ is **trace-class** if*

$$\sum_n \langle |T|\xi_n, \xi_n \rangle < \infty.$$

*We denote by $\mathcal{T}(\mathcal{H})$ the set of all trace-class operators on $\mathcal{H}$. The operator $T$ is **Hilbert–Schmidt** if*

$$\sum_n \langle T^*T\xi_n, \xi_n \rangle < \infty$$

*and we denote the set of all Hilbert–Schmidt operators by $\mathcal{HS}(\mathcal{H})$.*

*The **trace** is the map* $\mathrm{Tr} : \mathcal{T}(\mathcal{H}) \to \mathbb{C}$ *given by*

$$\mathrm{Tr}(T) = \sum_n \langle T\xi_n, \xi_n\rangle. \tag{10.21}$$

The definition above would not make a lot of sense if it depended on the basis, but we will soon see that it does not (Proposition 10.7.5).

An intuition to keep in mind, though the analogy is not always applicable, is that one can roughly think of $\mathcal{K}(\mathcal{H})$, $\mathcal{T}(\mathcal{H})$, $\mathcal{HS}(\mathcal{H})$, and $\mathcal{B}(\mathcal{H})$, as non-commutative versions of $c_0$, $\ell^1(\mathbb{N})$, $\ell^2(\mathbb{N})$, and $\ell^\infty(\mathbb{N})$ respectively. We will explore this connection in a bit.

It is a routine exercise that in the finite dimensional case the formula for the trace does not depend on the basis chosen; this follows from the equality $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$ and the fact that two matrices representing the same operator under different orthonormal bases are related by $T = VSV^*$ with $V$ a a unitary. In finite dimension we also have $\mathcal{T}(\mathcal{H}) = \mathcal{HS}(\mathcal{H}) = \mathcal{B}(\mathcal{H})$. In infinite-dimension, we cannot expect to apply (10.21) blindly to an arbitrary operator. For starters, if $T = I$ is the identity, then $\mathrm{Tr}(T) = \infty$. But it gets worse; for the unilateral shift $S$ corresponding to $\{\xi_n\}$, we would have $\mathrm{Tr}(S) = \sum_n \langle \xi_{n+1}, \xi_n\rangle = 0$. But if instead we consider the orthonormal basis $\{\eta_n\}$, where

$$\eta_{2n-1} = \frac{\xi_n + \xi_{n+1}}{\sqrt{2}}, \qquad \eta_{2n} = \frac{\xi_n - \xi_{n+1}}{\sqrt{2}},$$

we obtain $\langle S\eta_{2n-1}, \eta_{2n-1}\rangle = \frac{1}{2}$ and $\langle S\eta_{2n}, \eta_{2n}\rangle = -\frac{1}{2}$, showing that in general the series in (10.21) depends on the basis chosen, and may fail to exist for some bases and not others. It is tempting to say that each $\frac{1}{2}$ cancels with its corresponding $-\frac{1}{2}$, but the series does not converge regardless; and by reordering the basis we would be able to attain many different values (and the order of the basis can have no bearing in an expression defined in terms of the operator). This is the reason we restrict the domain of Tr to $\mathcal{T}(\mathcal{H})$. We summarize the main properties of $\mathcal{T}(\mathcal{H})$ in the next few results.

We first check that (10.21) works for all $T \in \mathcal{T}(\mathcal{H})$.

***10.7.2. Lemma.***

*Let* $T \in \mathcal{T}(\mathcal{H})$. *Then*

$$\sum_{k=1}^\infty |\langle T\xi_k, \xi_k\rangle| \leq \sum_{k=1}^\infty \langle |T|\xi_k, \xi_k\rangle < \infty \tag{10.22}$$

*In particular, the series* $\mathrm{Tr}(T) = \sum_k \langle T\xi_k, \xi_k\rangle$ *converges absolutely and*

$$|\mathrm{Tr}(T)| \leq \mathrm{Tr}(|T|).$$

*Proof.* Via the polar decomposition we write $T = V|T|$ with $V$ a partial isometry. Then, using Bessel's inequality (4.12),

$$\begin{aligned}
\sum_{k=1}^{n} |\langle T\xi_k, \xi_k\rangle| &= \sum_{k=1}^{n} |\langle V|T|\xi_k, \xi_k\rangle| = \sum_{k=1}^{n} |\langle |T|^{1/2}\xi_k, |T|^{1/2}V^*\xi_k\rangle| \\
&\leq \sum_{k=1}^{n} \| \, |T|^{1/2}\xi_k\| \, \| \, |T|^{1/2}V^*\xi_k\| \\
&\leq \left(\sum_{k=1}^{n} \| \, |T|^{1/2}\xi_k\|^2\right)^{1/2} \left(\sum_{k=1}^{n} \| \, |T|^{1/2}V^*\xi_k\|^2\right)^{1/2} \\
&= \left(\sum_{k=1}^{n} \langle |T|\xi_k, \xi_k\rangle\right)^{1/2} \left(\sum_{k=1}^{n}\sum_{j=1}^{\infty} |\langle |T|^{1/2}V^*\xi_k, \xi_j\rangle|^2\right)^{1/2} \\
&\leq \big(\operatorname{Tr}(|T|)\big)^{1/2} \left(\sum_{j=1}^{\infty}\sum_{k=1}^{n} |\langle \xi_k, V|T|^{1/2}\xi_j\rangle|^2\right)^{1/2} \\
&\leq \big(\operatorname{Tr}(|T|)\big)^{1/2} \left(\sum_{j=1}^{\infty} \|V|T|^{1/2}\xi_j\|^2\right)^{1/2} \\
&\leq \big(\operatorname{Tr}(|T|)\big)^{1/2} \left(\sum_{j=1}^{\infty} \||T|^{1/2}\xi_j\|^2\right)^{1/2} \\
&= \operatorname{Tr}(|T|).
\end{aligned}$$

As $n$ was arbitrary,

$$\sum_{k=1}^{\infty} |\langle T\xi_k, \xi_k\rangle| \leq \sum_{k=1}^{\infty} \langle |T|\xi_k, \xi_k\rangle < \infty. \qquad \square$$

The absolute convergence (10.22) of the series for $\operatorname{Tr}(T)$ should be expected in the sense that the trace is meant to be an invariant of the operator $T$, independent of presentation. This in particular means that a reorder of the orthonormal basis should not change the trace, and this requires and implies absolute convergence (see Exercises 9.8.18 and 9.8.19).

**10.7.3. Lemma.**

*Let $T \in \mathcal{T}(\mathcal{H})$ and $U \in \mathcal{B}(\mathcal{H})$ a unitary. Then $U^*TU \in \mathcal{T}(\mathcal{H})$. For any orthonormal basis $\{\eta_k\}$ we have*

$$\sum_{k} \langle |T|\eta_k, \eta_k\rangle = \operatorname{Tr}(|T|). \tag{10.23}$$

*In particular,*

$$\operatorname{Tr}(|T|) = \sum_k \sigma_k(T). \tag{10.24}$$

*Proof.* Using Parseval's equality and Tonelli (Corollary 2.7.17) to exchange the series,

$$\begin{aligned}\sum_n \langle |U^*TU|\xi_n, \xi_n\rangle &= \sum_n \langle U^*|T|U\xi_n, \xi_n\rangle = \sum_n \langle |T|^{1/2}U\xi_n, |T|^{1/2}U\xi_n\rangle \\ &= \sum_n \| \, |T|^{1/2}U\xi_n\|^2 = \sum_n \sum_k |\langle |T|^{1/2}U\xi_n, \xi_k\rangle|^2 \\ &= \sum_k \sum_n |\langle \xi_n, U^*|T|^{1/2}\xi_k\rangle|^2 = \sum_k \|U^*|T|^{1/2}\xi_k\|^2 \\ &= \sum_k \| \, |T|^{1/2}\xi_k\|^2 = \sum_k \langle |T|\xi_k, \xi_k\rangle < \infty.\end{aligned}$$

So $U^*TU \in \mathcal{T}(\mathcal{H})$ and $\operatorname{Tr}(U^*|T|U) = \operatorname{Tr}(|T|)$. When $\{\eta_k\}$ is an orthonormal basis there exists a unitary $U$ with $\eta_n = U\xi_n$, and this gives (10.23). Using an orthonormal basis of eigenvectors for $|T|$ (via the Spectral Theorem 10.6.12) we get (10.24). □

### *10.7.4. Lemma.*

*Let $T \in \mathcal{T}(\mathcal{H})$ be selfadjoint. Then $T^+, T^- \in \mathcal{T}(\mathcal{H})$, and* $\operatorname{Tr}(|T|) = \sum_k |\lambda_k(T)|$.

*Proof.* From Corollary 10.6.16 we have that $T^+ + T^- = |T|$. Then

$$\sum_k \langle T^+\xi_k, \xi_k\rangle \le \sum_k \langle |T|\xi_k, \xi_k\rangle < \infty.$$

So $T^+ \in \mathcal{T}(\mathcal{H})$, and similarly $T^- \in \mathcal{T}(\mathcal{H})$. The equality for the trace follows directly from (10.24) and the fact that $\sigma_k(T) = |\lambda_k(T)|$ (Exercise 10.6.17). □

### *10.7.5. Proposition.*

*With the notation above:*

*(i)* $\mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H}) \subset \mathcal{K}(\mathcal{H})$*;*

*(ii)* $T \in \mathcal{T}(\mathcal{H})$ *if and only if* $T^* \in \mathcal{T}(\mathcal{H})$*, and in that case* $\operatorname{Tr}(T^*) = \overline{\operatorname{Tr}(T)}$*;*

*(iii)* *if* $T \in \mathcal{T}(\mathcal{H})$ *and* $R, S \in \mathcal{B}(\mathcal{H})$, *then* $STR \in \mathcal{T}(\mathcal{H})$ *and*

$$|\operatorname{Tr}(STR)| \leq \operatorname{Tr}(|STR|) \leq \|S\|\,\|R\|\,\operatorname{Tr}(|T|); \tag{10.25}$$

*(iv)* $\mathcal{T}(\mathcal{H})$ *is a (proper and non-closed, in the infinite-dimensional case) selfadjoint ideal of* $\mathcal{B}(\mathcal{H})$*;*

*(v)* *if* $T \in \mathcal{T}(\mathcal{H})$ *and* $\{\xi_n\}$ *and* $\{\eta_n\}$ *are two orthonormal bases for* $\mathcal{H}$,

$$\operatorname{Tr}(T) = \sum_n \langle T\xi_n, \xi_n\rangle = \sum_n \langle T\eta_n, \eta_n\rangle; \tag{10.26}$$

*(vi)* *if* $T \in \mathcal{T}(\mathcal{H})$ *and* $S \in \mathcal{B}(\mathcal{H})$, *then* $\operatorname{Tr}(TS) = \operatorname{Tr}(ST)$*;*

*(vii)* *if* $\{\lambda_k\}$ *are the eigenvalues of* $T \in \mathcal{T}(\mathcal{H})$, *counting multiplicities, then*

$$\operatorname{Tr}(T) = \sum_k \lambda_k. \tag{10.27}$$

*Proof.*

(i) Assume first that $T \in \mathcal{T}(\mathcal{H})$ is positive. Let $P_n = \sum_{k=1}^n \langle \cdot, \xi_k\rangle\,\xi_k$ be the orthogonal projection onto the span of $\xi_1, \ldots, \xi_n$. Fix $\varepsilon > 0$. By hypothesis there exists $m$ such that $\sum_{k>m} \|T^{1/2}\xi_k\|^2 = \sum_{k>m}\langle T\xi_k, \xi_k\rangle < \varepsilon^2$. Then

$$\begin{aligned}
\|(T^{1/2} - T^{1/2}P_m)\xi\| = \|T^{1/2}(I - P_m)\xi\| &= \Big\|\sum_{k>m} \langle \xi, \xi_k\rangle\, T^{1/2}\xi_k\Big\| \\
&\leq \sum_{k>m} |\langle \xi, \xi_k\rangle|\,\|T^{1/2}\xi_k\| \\
&\leq \Big(\sum_{k>m} \|T^{1/2}\xi_k\|^2\Big)^{1/2} \Big(\sum_{k>m} |\langle \xi, \xi_k\rangle|^2\Big)^{1/2} \\
&\leq \varepsilon\,\|\xi\|.
\end{aligned}$$

Thus $\|T^{1/2} - T^{1/2}P_m\| \leq \varepsilon$, which shows that $T^{1/2}P_m \to T^{1/2}$. As $P_m$ is finite-rank, we have that $T^{1/2}P_m$ is finite-rank (Proposition 9.6.2). Then $T^{1/2}$ is a norm-limit of finite-rank operators, so compact. And then $T = (T^{1/2})^2 \in \mathcal{K}(\mathcal{H})$ since $\mathcal{K}(\mathcal{H})$ is an ideal.

For general $T \in \mathcal{T}(\mathcal{H})$ we have the polar decomposition $T = V|T|$; and $|T| \in \mathcal{T}(\mathcal{H})$ by definition of $\mathcal{T}(\mathcal{H})$. As $|T| \in \mathcal{K}(\mathcal{H})$ by the previous paragraph, $T \in \mathcal{K}(\mathcal{H})$ since $\mathcal{K}(\mathcal{H})$ is an ideal.

For the inclusion $\mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$, if $T$ is finite-rank then so is $|T|$. By the Spectral Theorem (10.6.12) we know that $|T|$ is a linear combination of rank-one projections. So $\sum_n \langle |T|\xi_n, \xi_n\rangle$ will be finite if $\sum_n \langle P\xi_n, \xi_n\rangle$ is finite for each rank-one projection $P$. Such a projection is of the form $P\xi = \langle \xi, \eta \rangle \, \eta$ for $\eta \in \mathcal{H}$ with $\|\eta\| = 1$. Then, using (4.14),

$$\sum_n \langle P\xi_n, \xi_n \rangle = \sum_n \langle \xi_n, \eta \rangle \langle \eta, \xi_n \rangle = \sum_n |\langle \eta, \xi_n \rangle|^2 = \|\eta\|^2 = 1,$$

and $T \in \mathcal{T}(\mathcal{H})$.

(ii) Knowing that both $|T|$ and $|T^*|$ have the same eigenvalues with the same multiplicities (by Corollaries 10.6.14 and 10.6.17) we obtain from (10.24)

$$\sum_k \langle |T|\xi_k, \xi_k \rangle = \sum_k \sigma_k(T) = \sum_k \sigma_k(T^*) = \sum_k \langle |T^*|\xi_k, \xi_k \rangle.$$

Hence $T^* \in \mathcal{T}(\mathcal{H})$. Now, since the series $\operatorname{Tr}(T)$ and $\operatorname{Tr}(T^*)$ converge by Lemma 10.7.2,

$$\operatorname{Tr}(T^*) = \sum_n \langle T^*\xi_n, \xi_n \rangle = \sum_n \langle \xi_n, T\xi_n \rangle = \sum_n \overline{\langle T\xi_n, \xi_n \rangle} = \overline{\operatorname{Tr}(T)}.$$

(iii) By Lemma 10.7.2 we have $|\operatorname{Tr}(STR)| \le \operatorname{Tr}(|STR|)$. From

$$T^*S^*ST \le \|S\|^2 \, T^*T$$

and the fact that the square root is operator-monotone (Proposition 10.4.6),

$$|ST| = (T^*S^*ST)^{1/2} \le \|S\| \, |T|.$$

Then

$$\operatorname{Tr}(|ST|) = \sum_n \langle |ST|\xi_n, \xi_n \rangle \le \|S\| \sum_n \langle |T|\xi_n, \xi_n \rangle = \|S\| \operatorname{Tr}(|T|) < \infty.$$

Using this twice,

$$\begin{aligned} |\operatorname{Tr}(STR)| &\le \|S\| \operatorname{Tr}(|TR|) = \|S\| \operatorname{Tr}(|R^*T^*|) \\ &\le \|S\| \, \|R^*\| \operatorname{Tr}(|T^*|) = \|S\| \, \|R\| \operatorname{Tr}(|T|). \end{aligned}$$

(iv) Suppose that $T_1, T_2 \in \mathcal{T}(\mathcal{H})$. Let $T_1 + T_2 = V|T_1 + T_2|$ be the polar decomposition. By (iii), $V^*T_1 \in \mathcal{T}(\mathcal{H})$ and $V^*T_2 \in \mathcal{T}(\mathcal{H})$, and in turn $\operatorname{Tr}(V^*T_1)$ and $\operatorname{Tr}(V^*T_2)$ exist. Now

$$\begin{aligned} \sum_n \langle |T_1 + T_2|\xi_n, \xi_n \rangle &= \sum_n \langle V^*(T_1 + T_2)\xi_n, \xi_n \rangle \\ &= \sum_n \langle V^*T_1\xi_n, \xi_n \rangle + \sum_n \langle V^*T_2\xi_n, \xi_n \rangle \\ &= \operatorname{Tr}(V^*T_1) + \operatorname{Tr}(V^*T_2) < \infty. \end{aligned}$$

As scalar multiples of trace-class operators are trivially trace-class, we have shown that $\mathcal{T}(\mathcal{H})$ is a subspace of $\mathcal{B}(\mathcal{H})$, and a two-sided ideal by (iii). The fact that $\mathcal{T}(\mathcal{H})$ is a proper ideal in the infinite-dimensional case is left as an exercise (Exercise 10.7.3).

(v) We know that (10.26) holds for positive $T \in \mathcal{T}(\mathcal{H})$ by (10.23). For arbitrary $T \in \mathcal{T}(\mathcal{H})$, let $A = \operatorname{Re} T$, $B = \operatorname{Im} T$. We have $A, B \in \mathcal{T}(\mathcal{H})$ by (ii) and (iv), and $A^+, A^-, B^+, B^- \in \mathcal{T}(\mathcal{H})$ by Lemma 10.7.4. Then, using (10.23),

$$\begin{aligned}\sum_n \langle T\eta_n, \eta_n\rangle &= \sum_n \langle A^+\eta_n, \eta_n\rangle - \sum_n \langle A^-\eta_n, \eta_n\rangle \\ &\qquad + i\sum_n \langle B^+\eta_n, \eta_n\rangle - i\sum_n \langle B^-\eta_n, \eta_n\rangle \\ &= \sum_n \langle A^+\xi_n, \xi_n\rangle - \sum_n \langle A^-\xi_n, \xi_n\rangle \\ &\qquad + i\sum_n \langle B^+\xi_n, \xi_n\rangle - i\sum_n \langle B^-\xi_n, \xi_n\rangle \\ &= \sum_n \langle T\xi_n, \xi_n\rangle.\end{aligned}$$

(vi) If $T \in \mathcal{T}(\mathcal{H})$ and $U$ is a unitary we can apply (v) to the orthonormal bases $\{\xi_k\}$ and $\{U\xi_j\}$ to get that $\operatorname{Tr}(U^*TU) = \operatorname{Tr}(T)$. Replacing $T$ with $UT \in \mathcal{T}(\mathcal{H})$ in this latter equality we get $\operatorname{Tr}(UT) = \operatorname{Tr}(TU)$. As any $S \in \mathcal{B}(\mathcal{H})$ can be written as a linear combination of four unitaries (see Remark 10.4.8), it follows that $\operatorname{Tr}(ST) = \operatorname{Tr}(TS)$.

(vii) If $T \in \mathcal{T}(\mathcal{H})$ is positive, we showed in Lemma 10.7.3 that $\operatorname{Tr}(T) = \sum_k \lambda_k$. For general $T$, the proof is far from trivial; it is known as **Lidskii's Theorem**, and we have included its details in Chapter C.

□

The inclusions $\mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H}) \subset \mathcal{K}(\mathcal{H})$ are proper when $\dim \mathcal{H} = \infty$ (Exercise 10.7.3). When both $ST$ and $TS$ are trace-class, we do have $\operatorname{Tr}(TS) = \operatorname{Tr}(ST)$, regardless of whether $S$ and $T$ are trace-class. Indeed, this follows from (10.27) and the fact that $TS$ and $ST$ have the same nonzero eigenvalues and multiplicities (Proposition 9.2.15 and Exercise 10.7.8; see Remark 10.7.6). But we cannot expect (10.26) above to hold in general when $T$ is not trace-class, even if one of $TS$ and $ST$ is trace-class. For instance fix an orthonormal basis $\{\eta_j\}$ and consider the associated matrix units as in Exercise 10.6.25 . Consider

$$S = \sum_k E_{2k,2k}, \qquad T = \sum_k \frac{1}{k^2}\, E_{2k,2k-1} + \sum_k E_{2k-1,2k}.$$

Then

$$ST = \sum_k \frac{1}{k^2}\, E_{2k,2k-1} \in \mathcal{T}(\mathcal{H}),$$

since $|ST| = \sum_k \frac{1}{k^2} E_{2k-1,2k-1}$ and thus $\mathrm{Tr}(|ST|) = \sum_k \frac{1}{k^2} < \infty$. With $TS$, on the other hand, we have

$$TS = \sum_k E_{2k-1,2k}, \tag{10.28}$$

and $|TS| = \sum_k E_{2k,2k}$, giving us $\mathrm{Tr}(|TS|) = \infty$. When an operator is not trace-class, the trace is not well-defined. The operator $TS$ from above is a good candidate to see that. We have

$$\sum_k \langle TS\eta_k, \eta_k\rangle = 0.$$

If we use the orthonormal basis $\{\xi_k\}$, where $\xi_{2k-1} = \frac{\eta_{2k-1}+\eta_{2k}}{\sqrt{2}}$, $\xi_{2k} = \frac{\eta_{2k-1}-\eta_{2k}}{\sqrt{2}}$, then

$$\sum_k \langle TS\xi_k, \xi_k\rangle = \frac{1}{2} - \frac{1}{2} + \frac{1}{2} - \frac{1}{2} + \cdots$$

If instead we form $\xi'_{2k-1} = \frac{\eta_{2k-1}+\sqrt{3}\,\eta_{2k}}{2}$, $\xi'_{2k} = \frac{\sqrt{3}\,\eta_{2k-1}-\eta_{2k}}{2}$, then

$$\sum_k \langle TS\xi'_k, \xi'_k\rangle = \frac{\sqrt{3}}{4} - \frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{4} - \frac{\sqrt{3}}{4} + \cdots$$

With a bit more effort one can show that given any $z_1, \ldots, z_n \in \mathbb{C}$ with $|z_j| \le \frac{1}{2}$ for all $j$, an orthonormal basis can be constructed such that the diagonal of $TS$ in said basis begins with $z_1, \ldots, z_n$ (Exercise 10.7.12). In the case of multiplying by the adjoint, we have that $S^*S \in \mathcal{T}(\mathcal{H})$ if and only if $SS^* \in \mathcal{T}(\mathcal{H})$ (Exercise 10.7.9).

*10.7.6. Remark*. Besides the fact that we deferred the proof for the equality (10.27) to Chapter C, it has not been established what "counting multiplicities" means for non-normal $T \in \mathcal{T}(\mathcal{H})$. A comparison with the finite-dimensional case shows that one counts multiplicities according to the existence of a basis of *generalized eigenvectors*. And that is what works here: the multiplicity of $\lambda \in \sigma(T)$ is defined to be

$$\alpha_\lambda(T) = \dim \bigcup_{k=1}^{\infty} \ker(T - \lambda I)^k.$$

It is worth noting that $\dim\ker(T - \lambda I)^n < \infty$ for any $T$ compact, $\lambda \in \mathbb{C} \setminus \{0\}$, and $n \in \mathbb{N}$ (Exercise 10.7.6). Also, $\ker(T - \lambda I)^n \subset \ker(T - \lambda I)^{n+1}$, and if $\ker(T-\lambda I)^n = \ker(T-\lambda I)^{n+1}$ then $\ker(T-\lambda I)^{n+k} = \ker(T-\lambda I)^n$ for all $k \in \mathbb{N}$ (Exercise 10.7.7).

The multiplicity $\alpha_\lambda(T)$ is well-defined for any $T \in \mathcal{K}(\mathcal{H})$, in the sense that it is always finite for $\lambda \in \sigma(T) \setminus \{0\}$:

***10.7.7. Proposition.***

*Let $T \in \mathcal{K}(\mathcal{H})$ and $\lambda \in \sigma(T) \setminus \{0\}$. Then $\alpha_\lambda(T) < \infty$.*

*Proof.* By Exercises 10.7.6 and 10.7.7, either $\alpha_\lambda(T) < \infty$ or all the inclusions $\ker(T-\lambda I)^n = \ker(T-\lambda I)^{n+1}$ are proper. So if $\alpha_\lambda(T) = \infty$ there exist unit vectors $\eta_1, \eta_2, \ldots$ such that $\eta_k \in \ker(T-\lambda I)^k \setminus \ker(T-\lambda I)^{k-1}$; since all these kernels are closed subspaces, we can choose the $\eta_k$ so that they are orthonormal. Then $\eta_k \perp (T-\lambda I)\eta_k$ and

$$\|T\eta_k\|^2 = \|(T-\lambda I)\eta_k + \lambda\eta_k\|^2 = \|(T-\lambda I)\eta_k\|^2 + |\lambda|^2 \geq |\lambda|^2. \tag{10.29}$$

As $\{\eta_k\}$ is orthonormal, it converges weakly to zero (for instance using Bessel's Inequality (4.12)). And $T$ is compact, so completely continuous (Proposition 9.6.5); therefore $\{T\eta_k\}$ converges in norm to 0. But then (10.29) forces $\lambda = 0$, a contradiction. □

**10.7.8. Definition.**

*The **trace-norm** on $\mathcal{T}(\mathcal{H})$ is given by $\|T\|_1 = \operatorname{Tr}(|T|)$. More generally, for $p \in [1, \infty)$ the $p^{\text{th}}$**-Schatten norm** is defined—for those $T$ where it makes sense—as $\|T\|_p = \operatorname{Tr}(|T|^p)^{1/p}$.*

In analogy with $\mathcal{T}(\mathcal{H})$, we can then define ideals

$$\mathcal{S}_p = \{T \in \mathcal{B}(\mathcal{H}) : \operatorname{Tr}(|T|^p) < \infty\},$$

the **Schatten ideals**, and a $p$-norm $\|T\|_p = \operatorname{Tr}(|T|^p)^{1/p}$ on $\mathcal{S}_p$. In particular, $\mathcal{S}_1 = \mathcal{T}(\mathcal{H})$, and $\mathcal{S}_2$, the **Hilbert–Schmidt** operators, becomes a Hilbert space with the inner product $\langle S, T\rangle = \operatorname{Tr}(T^*S)$. We have not yet proven that $\|\cdot\|_1$ is a norm, though! In this context, (10.25), when $R = I$, is Hölder's 1-$\infty$ inequality.

***10.7.9. Proposition.***

*The trace-norm is a norm. The space $(\mathcal{T}(\mathcal{H}), \|\cdot\|_1)$ is complete, and $\mathcal{F}(\mathcal{H})$ is dense in it. The trace is a bounded, positive, linear functional on $\mathcal{T}(\mathcal{H})$, with $\|\operatorname{Tr}\| = 1$.*

*Proof.* It is straightforward to see that $\operatorname{Tr}(|T|) = 0$ implies $T = 0$, and that $\operatorname{Tr}(|\lambda T|) = |\lambda| \operatorname{Tr}(|T|)$. So the only thing remaining to show that $\|\cdot\|_1$ is a norm, is that it satisfies the triangle inequality. But this is not too hard with the tools we got in Proposition 10.7.5. Indeed, if $S, T \in \mathcal{T}(\mathcal{H})$ and

$S+T = V|S+T|$ is the polar decomposition,

$$\begin{aligned}\|S+T\|_1 &= \operatorname{Tr}(|S+T|) = \operatorname{Tr}(V^*(S+T)) \\ &= \operatorname{Tr}(V^*S) + \operatorname{Tr}(V^*T) \le |\operatorname{Tr}(V^*S)| + |\operatorname{Tr}(V^*T)| \\ &\le \|V^*\| \operatorname{Tr}(|S|) + \|V^*\| \operatorname{Tr}(|T|) = \operatorname{Tr}(|S|) + \operatorname{Tr}(|T|) \\ &= \|S\|_1 + \|T\|_1.\end{aligned}$$

The completeness is a consequence of Theorem 10.7.11 below, together with Proposition 5.5.8; alternatively, one could recreate a proof modelled on the proof of Theorem 2.8.12. The linearity of the trace is a direct consequence of linearity of limits. If $S \ge 0$ then $\operatorname{Tr}(S) = \sum_n \langle S\xi_n, \xi_n\rangle \ge 0$. Also, from (10.22),

$$|\operatorname{Tr}(S)| \le \operatorname{Tr}(|S|) = \|S\|_1,$$

so $\|\operatorname{Tr}\| \le 1$. If $P$ is a rank-one orthogonal projection, then $\|P\|_1 = 1$ and $\operatorname{Tr}(P) = 1$, so $\|\operatorname{Tr}\| = 1$.

As for the density of $\mathcal{F}(\mathcal{H})$ in $\mathcal{T}(\mathcal{H})$, let $P_n$ denote the orthogonal projection onto $\operatorname{span}\{\xi_1, \dots, \xi_n\}$. Fix $T \in \mathcal{T}(\mathcal{H})$, and let $T_n = TP_n$; then $T_n \in \mathcal{F}(\mathcal{H})$. Via the Polar Decomposition we can write $|T - T_n| = U^*(T - T_n) = U^*T(I - P_n)$ for an appropriate partial isometry $U$. With $S = U^*T$,

$$\|T - T_n\|_1 = \operatorname{Tr}(|T - T_n|) = \operatorname{Tr}(S(I - P_n))$$

The operator $S$ is trace-class by Proposition 10.7.5. So

$$\|T_n - T\|_1 = \operatorname{Tr}(S(I - P_n)) = \sum_j \langle S(I - P_n)\xi_j, \xi_j\rangle = \sum_{j>n} \langle S\xi_j, \xi_j\rangle \xrightarrow[n\to\infty]{} 0$$

since the series for $\operatorname{Tr}(S)$ converges. So $\lim_{n\to\infty} \|T_n - T\|_1 = 0$. □

The next lemma should bring a sense of deja-vu from Section 5.6 (concretely, Propositions 5.6.1 and 5.6.8)

**10.7.10. Lemma.**

*Let $S \in \mathcal{B}(\mathcal{H})$. Then*

$$\begin{aligned}\|S\|_1 &= \sup\{|\operatorname{Tr}(ST)| : \ T \in \mathcal{K}(\mathcal{H}),\ \|T\| \le 1\} \\ &= \sup\{|\operatorname{Tr}(ST)| : \ T \in \mathcal{F}(\mathcal{H}),\ \|T\| \le 1\}.\end{aligned} \tag{10.30}$$

*The equality holds even when one—and hence the other—side is infinite.*

*Proof.* Assume first that $S \in \mathcal{T}(\mathcal{H})$. If $\|T\| \le 1$, then

$$|\operatorname{Tr}(ST)| \le \|T\| \operatorname{Tr}(|S|) \le \|S\|_1.$$

With $S = V|S|$ the polar decomposition, $\|S\|_1 = \operatorname{Tr}(|S|) = \operatorname{Tr}(V^*S) = \operatorname{Tr}(SV^*)$, which shows that (10.30) holds if we use $\mathcal{B}(\mathcal{H})$ instead of $\mathcal{K}(\mathcal{H})$ on the right. Fix an orthonormal basis $\{\xi_n\}$ and let $P_n$ be the orthogonal projection onto $\operatorname{span}\{\xi_1,\ldots,\xi_n\}$. Then $P_nV^*\in\mathcal{F}(\mathcal{H})\subset\mathcal{T}(\mathcal{H})$, with $\|P_nV^*\|\leq 1$ and

$$\begin{aligned}|\operatorname{Tr}(SV^*)-\operatorname{Tr}(SP_nV^*)| &= |\operatorname{Tr}(V^*S)-\operatorname{Tr}(P_nV^*SP_n)|\\ &=\left|\sum_{k>n}\langle V^*S\xi_n,\xi_n\rangle\right|=\sum_{k>n}\langle|S|\xi_n,\xi_n\rangle\end{aligned}$$

can be made arbitrarily small for $n$ big enough. This establishes (10.30).

When $S\notin\mathcal{T}(\mathcal{H})$, then $\sum_k\langle|S|\xi_k,\xi_k\rangle=\infty$. Choosing $T=P_n$, the orthogonal projection onto $\operatorname{span}\{\xi_1,\ldots,\xi_n\}$, then $T$ is finite-rank and $\|T\|=1$; and

$$\operatorname{Tr}(ST)=\operatorname{Tr}(SP_n)=\sum_{k=1}^n\langle|S|\xi_k,\xi_k\rangle.$$

So the right-hand-side in (10.30) is infinite, too. $\square$

Now we come to the main theorem of this section.

***10.7.11. Theorem. (Canonical Duality in $\mathcal{B}(\mathcal{H})$)***

*There is a canonical duality $\mathcal{K}(\mathcal{H})^*=\mathcal{T}(\mathcal{H})$ and $\mathcal{T}(\mathcal{H})^*=\mathcal{B}(\mathcal{H})$, where both equalities mean isometric isomorphism and the duality is given by $\langle T,S\rangle=\operatorname{Tr}(TS)$.*

*Proof.* We define $\Gamma:\mathcal{T}(\mathcal{H})\to\mathcal{K}(\mathcal{H})^*$ by $\Gamma(S)(T)=\operatorname{Tr}(TS)$. This map is clearly linear. The equality (10.30) shows that $\|\Gamma(S)\|=\|S\|_1$, so $\Gamma$ is isometric; in particular, $\Gamma(S)\in\mathcal{K}(\mathcal{H})^*$, and $\Gamma$ is injective. It remains to show that $\Gamma$ is surjective. For this, let $\varphi\in\mathcal{K}(\mathcal{H})^*$. Consider the form $[\xi,\eta]=\varphi(\xi\eta^*)$. The linearity of $\varphi$ makes this form sesquilinear. Also,

$$|[\xi,\eta]|=|\varphi(\xi\eta^*)|\leq\|\varphi\|\,\|\xi\eta^*\|=\|\varphi\|\,\|\xi\|\,\|\eta\|.$$

So the form is bounded, and by Proposition 10.1.5 there exists $S\in\mathcal{B}(\mathcal{H})$ with $\varphi(\xi\eta^*)=\langle S\xi,\eta\rangle$, for all $\xi,\eta\in\mathcal{H}$. Given $T\in\mathcal{K}(\mathcal{H})$ positive, by the Spectral Theorem we have

$$T=\sum_{j=1}^\infty\lambda_j(T)\,\xi_j\xi_j^*$$

for an orthonormal basis $\{\xi_j\}$ of eigenvectors for $T$. As $\varphi$ is bounded,

$$\varphi(T)=\sum_{j=1}^\infty\lambda_j(T)\,\varphi(\xi_j\xi_j^*)=\sum_{j=1}^\infty\lambda_j(T)\,\langle S\xi_j,\xi_j\rangle=\sum_{j=1}^\infty\langle S\xi_j,\lambda_j(T)\xi_j\rangle$$

$$= \sum_{j=1}^{\infty} \langle S\xi_j, T\xi_j \rangle = \sum_{j=1}^{\infty} \langle TS\xi_j, \xi_j \rangle = \operatorname{Tr}(TS).$$

Any $T \in \mathcal{K}(\mathcal{H})$ is a linear combination of four positive compact operators by Corollary 10.6.16; hence the equality $\varphi(T) = \operatorname{Tr}(TS)$ holds for all $T \in \mathcal{K}(\mathcal{H})$. By (10.30),

$$\|S\|_1 = \sup\{|\varphi(T)| : T \in \mathcal{K}(\mathcal{H}), \|T\| \leq 1\} = \|\varphi\|,$$

so $S \in \mathcal{T}(\mathcal{H})$.

The duality $\mathcal{T}(\mathcal{H})^* = \mathcal{B}(\mathcal{H})$ is shown very similarly, and we leave the details to the reader (Exercise 10.7.4). ☐

One interesting consequence of Theorem 10.7.11 is that $\mathcal{K}(\mathcal{H})^{**} = \mathcal{B}(\mathcal{H})$. The analogy of $\mathcal{K}(\mathcal{H})$, $\mathcal{T}(\mathcal{H})$, $\mathcal{B}(\mathcal{H})$ with $c_0$, $\ell^1(\mathbb{N})$, $\ell^\infty(\mathbb{N})$ should be clearer now. Another consequence, to be explored in Chapters 12 and 14, is that $\mathcal{B}(\mathcal{H})$ has a weak*-topology.

We finish the section with a characterization of trace-class operators that will be of use later. The full argument requires more advanced material (namely, the Spectral Theorem 12.4.4). The proof can be followed, though, if one assumes that $T$ is compact.

**_10.7.12. Proposition._**

*Let $T \in \mathcal{B}(\mathcal{H})$ such that $\sum_n \langle T\xi_n, \xi_n \rangle$ converges for any orthonormal basis $\{\xi_n\}$. Then $T \in \mathcal{T}(\mathcal{H})$.*

*Proof.* Because any reordering of an orthonormal basis is an orthonormal basis, the hypothesis is equivalent to requiring absolute convergence (see Exercises 9.8.18 and 9.8.19). Also, since

$$\operatorname{Re} \sum_n \langle T\xi_n, \xi_n \rangle = \sum_n \operatorname{Re} \langle T\xi_n, \xi_n \rangle = \sum_n \langle \operatorname{Re} T\xi_n, \xi_n \rangle$$

and similarly for the imaginary parts, the operators $\operatorname{Re} T$ and $\operatorname{Im} T$ also satisfy the hypothesis. As linear combinations of trace-class are trace-class, we may assume without loss of generality that $T$ is selfadjoint.

If $T$ is not compact, by Exercise 12.4.5 there exists $\lambda \in \sigma(T) \setminus \{0\}$ and $\delta > 0$, with either $\lambda - \delta > 0$ or $\lambda + \delta < 0$, such that the spectral projection $P = \mu_T(\lambda - \delta, \lambda + \delta)$ is infinite. With $\{\xi_n\}$ an orthonormal basis for the range of $P$ we have, with $\varepsilon = \min\{|\lambda - \delta|, |\lambda + \delta|\}$,

$$\sum_n |\langle T\xi_n, \xi_n \rangle| \geq \sum_n \varepsilon \langle \xi_n, \xi_n \rangle = \sum_n \varepsilon = \infty,$$

contradicting the hypothesis (any subseries of an absolutely-convergent series must converge; this cannot be said of arbitrary series). So $T$ is compact. Now the Spectral Theorem for Compact Operators (Theorem 10.6.12) allows us to write

$$T = \sum_n \lambda_n P_n,$$

where $\{\lambda_n\} \subset \mathbb{R}\setminus\{0\}$ is a sequence that converges to 0, and $\{P_n\}$ are pairwise orthogonal rank-one projections. If $T \notin \mathcal{T}(\mathcal{H})$, this means that

$$\sum_n |\lambda_n| = \infty.$$

Let $\{\xi_n\}$ be unit vectors with $P_n\xi_n = \xi_n$. Then $\{\xi_n\}$ is orthonormal and

$$\sum_n |\langle T\xi_n, \xi_n\rangle| = \sum_n |\lambda_n| = \infty,$$

a contradiction. Thus $T \in \mathcal{T}(\mathcal{H})$. $\square$

**10.7.1. Exercises.**

(10.7.1) Proving the inequality $|\operatorname{Tr}(T)| \leq \operatorname{Tr}(|T|)$ in (iii) of Proposition 10.7.5, it is tempting to get a direct proof by using

$$|\langle T\xi_n, \xi_n\rangle| \leq \langle |T|\xi_n, \xi_n\rangle,$$

from where the inequality would follow directly. Show that this inequality does not hold in general.

(10.7.2) Let $\varphi : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ be linear and *positive*, that is $\varphi(T) \geq 0$ if $T \geq 0$. Prove the Cauchy–Schwarz inequality

$$|\varphi(S^*T)| \leq \varphi(S^*S)^{1/2}\varphi(T^*T)^{1/2}.$$

(10.7.3) Show that if $\dim \mathcal{H} = \infty$ then the inclusions $\mathcal{F}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H}) \subset \mathcal{K}(\mathcal{H})$ are proper.

(10.7.4) Show that $\mathcal{T}(\mathcal{H})^* = \mathcal{B}(\mathcal{H})$; that is, show that there is an isometric isomorphism $\Gamma : \mathcal{B}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})^*$ such that $\Gamma(T)(S) = \operatorname{Tr}(TS)$.

(10.7.5) Let $T \in \mathcal{B}(\mathcal{H})$. Show that if $\lambda \neq 0$ then $\ker T \cap \ker(T - \lambda I)^n = \{0\}$ for all $n$.

(10.7.6) Show that $\dim \ker(T - \lambda I)^n < \infty$ for any $T$ compact, $\lambda \in \mathbb{C} \setminus \{0\}$, and $n \in \mathbb{N}$.

(10.7.7) Show that if $\ker(T - \lambda I)^n = \ker(T - \lambda I)^{n+1}$ then $\ker(T - \lambda I)^{n+k} = \ker(T - \lambda I)^n$ for all $k \in \mathbb{N}$.

(10.7.8) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $T : \mathcal{H} \to \mathcal{K}$, $S : \mathcal{K} \to \mathcal{H}$ linear and such that $TS \in \mathcal{T}(\mathcal{H})$, $ST \in \mathcal{T}(\mathcal{K})$. Show that $ST$ and $TS$ have the same nonzero eigenvalues, with $\alpha_\lambda(ST) = \alpha_\lambda(TS)$ for all $\lambda \in \sigma(ST) \setminus \{0\}$. Conclude that $\operatorname{Tr}(TS) = \operatorname{Tr}(ST)$.

(10.7.9) Let $S \in \mathcal{B}(\mathcal{H})$. Show that $S^*S \in \mathcal{T}(\mathcal{H})$ if and only if $SS^* \in \mathcal{T}(\mathcal{H})$.

(10.7.10) Let $S \in \mathcal{T}(\mathcal{H})$. Show that

$$\inf\{\|S+R\|_1 : \ \mathrm{Tr}(R)=0\} = |\,\mathrm{Tr}(S)|. \tag{10.31}$$

(10.7.11) Let $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}$ such that

$$\sum_n \|\xi_n\|^2 < \infty, \qquad\qquad \sum_n \|\eta_n\|^2 < \infty.$$

Use Proposition 10.7.12 to show that the operator $S = \sum_n \xi_n\eta_n^*$ is trace-class.

(10.7.12) Let $TS$ be as in (10.28) and let $z_1, \ldots, z_n \in \mathbb{C}$ with $|z_j| \leq \frac{1}{2}$ for all $j$. Show that there exists an orthonormal basis $\{\xi_k\}$ such that the diagonal of $TS$ in said basis begins with $z_1, \ldots, z_n$.

(10.7.13) Let $\xi, \eta \in \mathcal{H}$. Show that

$$\|\xi\eta^*\|_1 = \|\xi\|\,\|\eta\|. \tag{10.32}$$

(10.7.14) Let $\xi, \eta \in \mathcal{H}$ be unit vectors and $P = \xi\xi^*$, $Q = \eta\eta^*$ the corresponding rank-one projections. Show that

$$\|P-Q\|_1 = 2\sqrt{1-|\langle \xi, \eta\rangle|^2}.$$

(10.7.15) Let $\mathcal{H}$ be a Hilbert space. Show that the Banach space $\mathcal{T}(\mathcal{H})$ is separable if and only if $\mathcal{H}$ is separable.

(10.7.16) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $U : \mathcal{H} \to \mathcal{K}$ a unitary. Show that $A \in \mathcal{T}(\mathcal{H})$ if and only if $UAU^* \in \mathcal{T}(\mathcal{K})$.

(10.7.17) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $U : \mathcal{H} \to \mathcal{K}$ a unitary. Show that $\mathrm{Tr}(UAU^*) = \mathrm{Tr}(A)$ for all $A \in \mathcal{T}(\mathcal{H})$ *(this is slightly less trivial than it looks, since we are using the trace in two different spaces).*

CHAPTER

# 11

# C*-Algebras

There have been several cases in the text so far, where the manipulations seem to go through algebraically regardless of what the object actually is. This happened both with spaces of functions in Section 7.4 and with operators in Chapter 10. We will now see that there is a unifying theme, and that is the notion of C*-algebra. We will have to first develop the theory for quite a bit, but it will eventually become clear that the axioms in the definition of C*-algebra correctly encompass both what happens in $C(X)$ and what happens in $\mathcal{B}(\mathcal{H})$.

This chapter contains what one could call the basic theory of C*-algebras. Later, with more results about operators at hand, we will extend our discussion in Chapter 13.

## 11.1. C*-Algebra Basics

**11.1.1. Definition.**

*A **C*-algebra** is a nonzero Banach algebra $\mathcal{A}$ with the following additional structure: given $a, b \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{C}$,*

- *an **involution** $a \longmapsto a^*$; that is, a map $* : \mathcal{A} \to \mathcal{A}$*

  *(i) $(\alpha a + \beta b)^* = \bar{\alpha} a^* + \bar{\beta} b^*$;*

> *(ii)* $(ab)^* = b^* a^*$;
>
> *(iii)* $a^{**} = a$.
>
> - *The **C*-identity**: for all $a \in \mathcal{A}$, $\|a\|^2 = \|a^* a\|$.*
>
> *A **$*$-algebra** is an algebra with an involution.*
>
> *A **C*-norm** on a $*$-algebra is a submultiplicative norm that satisfies the C*-identity.*
>
> *$\mathcal{A}$ is **unital** if it has a multiplicative identity $I_{\mathcal{A}}$.*
>
> *If $\mathcal{A}$ is unital, an element $u \in \mathcal{A}$ is **unitary** if $uu^* = uu^* = I_{\mathcal{A}}$.*

The nonzero requirement is to avoid the trivial case $\mathcal{A} = \{0\}$.

When the C*-algebra $\mathcal{A}$ us unital, its identity is most often denoted as $I$ or $I_{\mathcal{A}}$; in some cases, depending on context, it might be denoted as 1. It is trivial to check that $I_{\mathcal{A}}^*$ is also an identity, so $I_{\mathcal{A}}^* = I_{\mathcal{A}}$. Then $\|I_{\mathcal{A}}\|^2 = \|I_{\mathcal{A}}^* I_{\mathcal{A}}\| = \|I_{\mathcal{A}}^2\| = \|I_{\mathcal{A}}\|$, so $\|I_{\mathcal{A}}\| = 1$; properly, we can also have $\|I_{\mathcal{A}}\| = 0$, but this forces $\mathcal{A} = \{0\}$, which is not too interesting as a C*-algebra.

We also have $\|a\|^2 = \|a^* a\| \leq \|a^*\| \|a\|$, so for any $a$ we have $\|a\| \leq \|a^*\|$. As we can also do this for $a^*$, we get that $\|a\| = \|a^*\|$.

A simple consequence of the C*-identity is that in a C*-algebra, $a = 0$ if and only if $a^* = 0$ if and only if $a^* a = 0$ if and only if $aa^* = 0$. This follows directly from $\|a\|^2 = \|a^* a\| = \|a^*\|$.

*11.1.2. Examples*. Although the definition seems to have a lot of ingredients, several C*-algebras of different flavours are already familiar to the reader. For instance,

(i) $C_0(T)$ with $T$ locally compact Hausdorff and the uniform norm, where the involution is given by $f^*(t) = \overline{f(t)}$;

(ii) the above example includes for instance $c_0 = C_0(\mathbb{N})$ and $\ell^\infty(\mathbb{N}) = C_b(\mathbb{N})$ (more information on this C*-algebra will be given in Example 11.2.9 );

(iii) $M_n(\mathbb{C})$, where the adjoint is the conjugate transpose, and with the operator norm;

(iv) $\mathcal{B}(\mathcal{H})$, the algebra of bounded operators on a Hilbert space, with the operator norm and the involution $T^*$ given by $\langle T^* x, y\rangle = \langle x, Ty\rangle$.

*11.1.3. Remark* . In some sense, the "usual" C*-algebras are the separable ones; non-separable C*-algebras often carry more structure (typically, they are von Neumann algebras). In the above examples, $c_0$ and $M_n(\mathbb{C})$ are separable, as is $C(T)$ when $T$ is compact and metrizable (Proposition 7.4.24). On the other hand, $\ell^\infty(\mathbb{N})$ is not separable, and neither is $\mathcal{B}(\mathcal{H})$ when $\mathcal{H}$ is infinite-dimensional.

When a C*-algebra is not unital, it is possible to enlarge it to contain a unit. This we already knew for Banach algebras from Proposition 9.2.21, but the unitization obtained there does not satisfy the C*-identity.

**11.1.4. Proposition. (Unitization)**

*Let $\mathcal{A}$ be a non-unital C*-algebra. Then there exists a unital C*-algebra $\tilde{\mathcal{A}}$ and an isometric monomorphism $\pi : \mathcal{A} \to \tilde{\mathcal{A}}$ such that*

*(i) $\pi(\mathcal{A})$ is an ideal in $\tilde{\mathcal{A}}$ of codimension 1;*

*(ii) if $\mathcal{A}$ is abelian, so is $\tilde{\mathcal{A}}$;*

*(iii) if $\mathcal{B}$ is a unital C*-algebra and $\rho : \mathcal{A} \to \mathcal{B}$ is a $*$-monomorphism, there exists a unique unital $*$-monomorphism $\tilde{\rho} : \tilde{\mathcal{A}} \to \mathcal{B}$ with $\tilde{\rho}|_{\mathcal{A}} = \rho$.*

*Proof.* We construct $\tilde{\mathcal{A}}$ as in Proposition 9.2.21. That is, we let $\tilde{\mathcal{A}} = \mathcal{A} \oplus \mathbb{C}$, and $\pi(a) = (a, 0)$. On $\tilde{\mathcal{A}}$ we consider pointwise addition and multiplication by scalars, the multiplication

$$(a, \lambda)(b, \mu) = (ab + \lambda b + \mu a, \lambda\mu),$$

and the involution

$$(a, \lambda)^* = (a^*, \bar{\lambda}).$$

Then $\tilde{\mathcal{A}}$ is a unital $*$-algebra with unit $(0, 1)$. As mentioned above, the problem is that the norm for a C*-algebra needs to satisfy more conditions. Define a norm by

$$\|(a, \lambda)\| = \sup\{\|ab + \lambda b\| : \ b \in \mathcal{A}, \ \|b\| \leq 1\}. \tag{11.1}$$

It is clear that this is homogeneous and subadditive. If $\|(a, \lambda)\| = 0$, we get $ab + \lambda b = 0$ for all $b \in \mathcal{A}$; if $\lambda = 0$, we get $ab = 0$ for all $b$, and in in particular, $aa^* = 0$, so $a = 0$. And if $\lambda \neq 0$, then $-a/\lambda$ is a unit for $\mathcal{A}$, a contradiction.

The norm is also submultiplicative (Exercise 11.1.2), and it is also a C*-norm:

$$\begin{aligned}
\|(a, \lambda)\|^2 &= \sup\{\|ab + \lambda b\|^2 : \ \|b\| \leq 1\} \\
&= \sup\{\|b^* a^* ab + \lambda\, b^* a^* b + \bar{\lambda}\, b^* ab + |\lambda|^2 b^* b\| : \ \|b\| \leq 1\} \\
&= \sup\{\|b^* \big(a^* ab + \lambda\, a^* b + \bar{\lambda}\, ab + |\lambda|^2 b\big)\| : \ \|b\| \leq 1\} \\
&\leq \sup\{\|a^* ab + \lambda\, a^* b + \bar{\lambda}\, ab + |\lambda|^2 b\| : \ \|b\| \leq 1\}
\end{aligned}$$

$$= \|(a^*a + 2\mathrm{Re}\,\bar{\lambda}a, |\lambda|^2)\| = \|(a,\lambda)^*(a,\lambda)\|$$
$$\leq \|(a,\lambda)^*\|\,\|(a,\lambda)\|.$$

Then $\|(a,\lambda)\| \leq \|(a,\lambda)^*\|$. Working with $(a^*,\bar{\lambda})$ we get the reverse inequality and then $\|(a,\lambda)\| = \|(a,\lambda)^*\|$; this makes the inequalities equalities and thus $\|(a,\lambda)^2\| = \|(a,\lambda)^*(a,\lambda)\|$. This new norm extends the norm of $\mathcal{A}$, for if $a \in \mathcal{A}$ then $\|\pi(a)\| = \|(a,0)\| = \sup\{\|ab\| : \; b \in \mathcal{A}, \; \|b\| \leq 1\}$. Taking $b = a^*/\|a\|$ shows that $\pi$ is isometric.

Next we check that $\tilde{\mathcal{A}}$ is complete. Let $\{(a_n,\lambda_n)\}$ be a Cauchy sequence in $\tilde{\mathcal{A}}$. We note that $\mathcal{A}$ is closed in $\tilde{\mathcal{A}}$, for we just showed that the norm of $\tilde{\mathcal{A}}$ agrees on $\mathcal{A}$ with the norm of $\mathcal{A}$. Using Hahn–Banach (Corollary 5.7.19), there exists $\varphi \in \tilde{\mathcal{A}}^*$ with $\varphi|_{(\mathcal{A},0)} = 0$ and $\varphi(0,\lambda) = \lambda$ for all $\lambda \in \mathbb{C}$. Then

$$|\lambda_n - \lambda_m| = |\varphi((a_n,\lambda_n) - (a_m,\lambda_m))| \leq \|\varphi\|\,\|(a_n,\lambda_n) - (a_m,\lambda_m)\|$$

and thus $\{\lambda_n\}$ is Cauchy and converges to some $\lambda \in \mathbb{C}$. Now

$$\|a_n - a_m\| = \|(a_n,0) - (a_m,0)\| = \|(a_n,\lambda_n) - (a_m,\lambda_m) - (0,\lambda_n) + (0,\lambda_m)\|$$
$$\leq \|(a_n,\lambda_n) - (a_m,\lambda_m)\| + \|(0,\lambda_n) - (0,\lambda_m)\|$$
$$= \|(a_n,\lambda_n) - (a_m,\lambda_m)\| + |\lambda_n - \lambda_m|$$

and $\{a_n\}$ is Cauchy; as $\mathcal{A}$ is complete, there exists $a \in \mathcal{A}$ with $a_n \to a$. Then

$$\|(a_n,\lambda_n) - (a,\lambda)\| \leq \|a_n - a\| + |\lambda_n - \lambda| \to 0$$

and the sequence $\{(a_n,\lambda_n)\}$ is convergent, showing that $\tilde{\mathcal{A}}$ is complete.

The embedding $\pi : a \longmapsto (a,0)$ is clearly a $*$-monomorphism, and

$$\|\pi(a)\| = \|(a,0)\| = \|a\|.$$

It is straightforward to check that $\tilde{\mathcal{A}}$ is abelian if $\mathcal{A}$ is. And, as in the proof of Proposition 9.2.21, $\pi(\mathcal{A})$ is a maximal ideal of codimension 1.

Finally, suppose that $\mathcal{B}$ is a unital C$^*$-algebra, with a $*$-monomorphism $\rho : \mathcal{A} \to \mathcal{B}$. Define $\tilde{\rho} : \tilde{\mathcal{A}} \to \mathcal{B}$ by

$$\tilde{\rho}(a,\lambda) = \rho(a) + \lambda I_{\mathcal{B}}.$$

From the fact that $\rho$ is a $*$-homomorphism, we obtain that so is $\tilde{\rho}$. The uniqueness is straightforward from the requirement that $\tilde{\rho}$ be unital and extend $\rho$. So it remains to see that $\tilde{\rho}$ is injective. If $\tilde{\rho}(a,\lambda) = 0$, this means that $\rho(a) + \lambda I_{\mathcal{B}} = 0$. If $\lambda \neq 0$, then $\rho(a') = I_{\mathcal{B}}$, where $a' = -a/\lambda$. For any $b \in \mathcal{B}$, $\rho(a'b) = \rho(a')\rho(b) = \rho(b)$; by the injectivity, $a'b = b$. We can similarly get that $ba' = b$ for all $b \in \mathcal{A}$, implying that $\mathcal{A}$ is unital, a contradiction. So $\lambda = 0$; then $\rho(a) = 0$ and so $a = 0$ by the injectivity of $\rho$, giving us $(a,\lambda) = (0,0)$. $\square$

*11.1.5. Remark.* The statement in Proposition 11.1.4 seems a bit unsatisfactory in that it requires $\mathcal{A}$ to be non-unital. What happens if we unitize a unital C*-algebra? No C*-protection agency will come to detain us, but we do run into a couple problems. The first one is that if $\mathcal{A}$ is unital, the unit of $\mathcal{A}$ will not be the unit in $\tilde{\mathcal{A}}$. The second one is that in the proof that $\|(a,\lambda)\| = \sup\{\|ab+\lambda b| : \|b\| \leq 1\}$ is a norm, we used that $\mathcal{A}$ was non-unital. And this was no gimmick of the proof: if we construct $\tilde{\mathcal{A}}$ when $\mathcal{A}$ is unital, we get $\|(I_{\mathcal{A}}, -1)\| = 0$, so we do not get a norm. There is a way to go around this, which is to define

$$\|(a,\lambda)\| = \max\big\{|\lambda|, \sup\{\|ab+\lambda b\| : \ \|b\| \leq 1\}\big\}.$$

One can show that this is a norm that makes $\tilde{\mathcal{A}}$ a C*-algebra, and that when $\mathcal{A}$ is not unital we recover $\tilde{\mathcal{A}}$ as in Proposition 11.1.4 (Exercise 11.1.4). One interesting consequence of this, from the fact that a C*-algebra admits a unique C*-norm, is the not entirely obvious inequality

$$|\lambda| \leq \sup\{\|ab+\lambda b| : \ \|b\| \leq 1\}, \qquad a \in \mathcal{A}, \tag{11.2}$$

that holds when $\mathcal{A}$ is non-unital, and fails when $\mathcal{A}$ is unital (Exercise 11.1.5).

There are some cases where the unitization can be made explicit. Recall the one-point compactification of a locally compact Hausdorff space from Proposition 1.8.27.

***11.1.6. Proposition.***

*Let $T$ be a locally compact Hausdorff space. Then $\widetilde{C_0(T)} \simeq C(T_\infty)$, where $T_\infty$ is the one-point compactification of $T$.*

*Proof.* Let $f \in C_0(T)$ and $\{t_j\} \subset T$ a net with $t_j \to \infty$ in $T_\infty$. Given $\varepsilon > 0$ there exists $K \subset T$ compact with $|f(t)| < \varepsilon$ for $t \notin K$. As $T_\infty \setminus K$ is an open neighbourhood of $\infty$, there exists $j_0$ such that $t_j \in T_\infty \setminus K$ for all $j \geq j_0$. It follows that $f(t_j) \to 0$, and so we can extend $f$ as $f(\infty) = 0$ and still be continuous. With $f$ extended as $\tilde{f}|_T = f$ and $\tilde{f}(\infty) = 0$, we define $\pi : \widetilde{C_0(T)} \to C(T_\infty)$ by

$$\pi(f,\lambda) = \tilde{f} + \lambda 1.$$

It is clear that $\pi$ is a $*$-homomorphism. If $\pi(f,\lambda) = 0$ then $\lambda = 0$ by evaluating at $\infty$, and then $\tilde{f} = 0$ which implies $f = 0$; so $\pi$ is injective. Given $g \in C(T_\infty)$, let $f = g|_T - g(\infty)$. Then $g = \pi(f, g(\infty))$ and $\pi$ is surjective. Soon, with more theory at our disposal, we will be able to say that $\pi$ being a bijective isomorphism is enough to conclude that it is isometric. For now, we need an ad-hoc argument:

$$\|(f,\lambda)\| = \sup\{\|fg+\lambda g\| : \ \|g\| \leq 1\} = \sup\{|f(t)g(t)+\lambda g(t)| : \ |g(t)| \leq 1\}$$

$$= \sup\{|f(t) + \lambda| : \ t \in T\} = \sup\{|\tilde{f}(t) + \lambda| : \ t \in T_\infty\}$$
$$= \|\tilde{f} + \lambda 1\| = \|\pi(f, \lambda)\|.$$

Thus $\widetilde{C_0(T)}$ and $C(T_\infty)$ are isometrically isomorphic as C*-algebras via a natural isomorphism. □

Since C*-algebras are Banach algebras, everything said in Section 9.2 applies. The adjoint is a new feature, though, and we have to account for that. Here is some information about the adjoint.

**11.1.7. Proposition.**

*Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Then*

*(i) if $\mathcal{A}$ is unital and $a \in \mathcal{A}$ is invertible, then $a^*$ is invertible and $(a^*)^{-1} = (a^{-1})^*$;*

*(ii) $\sigma(a^*) = \{\bar{\lambda} : \ \lambda \in \sigma(a)\}$.*

*Proof.*

(i) We have

$$(a^{-1})^* a^* = (aa^{-1})^* = I_{\mathcal{A}}^* = I_{\mathcal{A}}, \quad a^*(a^{-1})^* = (a^{-1}a)^* = I_{\mathcal{A}}^* = I_{\mathcal{A}}. \tag{11.3}$$

(ii) From above we have that $a - \bar{\lambda} I_{\mathcal{A}} = (a - \lambda I_{\mathcal{A}})^*$ is invertible if and only if $a - \lambda I_{\mathcal{A}}$ is invertible. Therefore $\sigma(a^*) = \{\bar{\lambda} : \ \lambda \in \sigma(a)\}$. □

**11.1.8. Definition.**

*Given a subset $S \subset \mathcal{A}$, the* ***C*-algebra generated by*** *$S$ is the smallest C*-subalgebra of $\mathcal{A}$ that contains $S$, that we denote by $C^*(S)$.*

The C*-subalgebra $C^*(S)$ exists for any subset $S \subset \mathcal{A}$, since intersections of C*-algebras are again C*-algebras; in particular, we will often consider the C*-subalgebra $C^*(a)$ of $\mathcal{A}$ generated by $a \in \mathcal{A}$ (Exercise 11.1.7). We have

$$C^*(a) = \overline{\{p(a, a^*) : \ p \ \text{ non-commutative polynomial with } p(0,0) = 0\}}$$

(Exercise 11.1.8). When $\mathcal{A}$ is unital and $a$ is not invertible, there is a small distinction to be made, as we may consider both the unital and non-unital subalgebras generated by $a$. When the distinction is possible and we don't mention it explicitly, we will assume that we take the non-unital version. It is worth mentioning that the "non-unital" subalgebra may be unital, just with a different unit. For an example consider the C*-subalgebra of $M_2(\mathbb{C})$ generated by $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

We now establish some basic properties of $*$-homomorphisms. We will complete the picture in Propositions 11.4.9 and 11.2.10. Eventually we will be able to show that the image of a $*$-homomorphism of C*-algebras is closed.

This next definition is a entirely similar to the one in $\mathcal{B}(\mathcal{H})$.

**11.1.9. Definition.**

*An element $a \in \mathcal{A}$ is **normal** if $a^*a = aa^*$.*

***11.1.10. Lemma.***

*Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$ normal. Then* $\operatorname{spr}(a) = \|a\|$.

*Proof.* A careful examination of the corresponding argument in the proof of Proposition 10.3.3 shows that only Theorem 9.2.16 and the C*-identity are used. □

As simple as it looks, a remarkable consequence of Lemma 11.1.10 is that there is a unique C*-norm on $\mathcal{A}$. Because for any $a \in \mathcal{A}$ we have

$$\|a\| = \|a^*a\|^{1/2} = \operatorname{spr}(a^*a)^{1/2}. \tag{11.4}$$

***11.1.11. Proposition.***

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then $\pi$ is bounded, and $\|\pi\| \leq 1$.*

*Proof.* Assume first that $\mathcal{A}$ is unital; since $\pi(I_{\mathcal{A}})$ is a unit for $\pi(\mathcal{A})$ we may replace $\mathcal{B}$ with $\overline{\pi(\mathcal{A})}$ and hence assume that $\pi(I_{\mathcal{A}}) = I_{\mathcal{B}}$. If $a \in GL(\mathcal{A})$, then $\pi(a) \in GL(\mathcal{B})$, since $ab = ba = I_{\mathcal{A}}$ implies $\pi(a)\pi(b) = \pi(b)\pi(a) = I_{\mathcal{B}}$. As a consequence, for any $a \in \mathcal{A}$ we have $\sigma(\pi(a)) \subset \sigma(a)$ (Exercise 9.2.14). By Lemma 11.1.10,

$$\|\pi(a)\|^2 = \|\pi(a^*a)\| = \operatorname{spr}(\pi(a^*a)) \leq \operatorname{spr}(a^*a) = \|a^*a\| = \|a\|^2.$$

So $\|\pi(a)\| \leq \|a\|$ for all $a \in \mathcal{A}$.

When $\mathcal{A}$ is not unital, we consider its unitization $\tilde{\mathcal{A}}$, and same with $\mathcal{B}$ if needed. Then we define $\tilde{\pi}(a, \lambda) = \pi(a) + \lambda I_{\mathcal{B}}$; it is straightforward to check that $\tilde{\pi}$ is still a $*$-homomorphism, $\|\pi\| \leq \|\tilde{\pi}\|$, and the proof above applies. □

There is a subtlety in the proof of Proposition 11.1.11, in that a priori the spectral radius depends on the algebra (as discussed before Proposition 9.2.8), so replacing $\mathcal{B}$ with a subalgebra is suspect. But Lemma 11.1.10 shows that the spectral radius of a normal element does not depend on the subalgebra. In fact,

***11.1.12. Proposition.***

*Let $a \in \mathcal{B} \subset \mathcal{A}$.*

*(i) If $\mathcal{A}, \mathcal{B}$ unital with $I_{\mathcal{B}} = I_{\mathcal{A}}$, then $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a)$.*

*(ii) If $\mathcal{B}$ unital with $I_{\mathcal{B}}$ not the unit of $\mathcal{A}$, then $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a) \cup \{0\}$.*

*(iii) If $\mathcal{B}$ is not unital, then $\sigma_{\mathcal{B}}(a) = \sigma_{\mathcal{A}}(a)$.*

*Proof.* We consider first the case where $I_{\mathcal{B}} = I_{\mathcal{A}}$. Assume that $a^* = a$. Then $\sigma_{\mathcal{A}}(a), \sigma_{\mathcal{B}}(a) \subset \mathbb{R}$, so they have no holes and thus are equal by Proposition 9.2.8.

For the general case, we want to show that $a \in GL(\mathcal{A})$ if and only if $a \in GL(\mathcal{B})$. It is obvious that if $a \in GL(\mathcal{B})$, then $a \in GL(\mathcal{A})$. Conversely, suppose that $a \in GL(\mathcal{A})$. That is, there exists $b \in \mathcal{A}$ with $ab = ba = I_{\mathcal{A}}$. Then $aa^*b^*b = a(ba)^*b = ab = I_{\mathcal{A}}$, and similarly $b^*baa^* = I_{\mathcal{A}}$; so $aa^* \in GL(\mathcal{A})$. By the first paragraph, $aa^* \in GL(\mathcal{B})$. So there exists $c \in \mathcal{B}$ with $aa^*c = I_{\mathcal{A}}$, and so $a$ has a right inverse on $\mathcal{B}$. Repeating the argument with $a^*$ instead of $a$, we get that $a^*$ has a right inverse, so $a$ has a left inverse in $\mathcal{B}$. Having both a right and a left inverse, $a \in GL(\mathcal{B})$.

Now, when $I_{\mathcal{B}}$ is not the unit of $\mathcal{A}$, if necessary we replace $\mathcal{A}$ with its unitization. As we have $I_{\mathcal{B}} \neq I_{\mathcal{A}}$, no element of $\mathcal{B}$ can be invertible in $\mathcal{A}$, for we would get $I_{\mathcal{A}} = b^{-1}b = b^{-1}bI_{\mathcal{B}} = I_{\mathcal{A}}I_{\mathcal{B}} = I_{\mathcal{B}}$, a contradiction. This means that $0 \in \sigma_{\mathcal{A}}(a)$. Also, if $\lambda \neq 0$ and $a - \lambda I_{\mathcal{B}}$ is invertible in $\mathcal{B}$, there exists $c \in \mathcal{B}$ with $c(a - \lambda I_{\mathcal{B}}) = (a - \lambda I_{\mathcal{B}}) = I_{\mathcal{B}}$. Let $b = c + \lambda^{-1}(I_{\mathcal{A}} - I_{\mathcal{B}})$. Then, as $(I_{\mathcal{A}} - I_{\mathcal{B}})a = 0$,

$$b(a - \lambda I_{\mathcal{A}}) = c(a - \lambda I_{\mathcal{B}}) + \lambda^{-1}\lambda(I_{\mathcal{A}} - I_{\mathcal{B}}) = I_{\mathcal{A}},$$

and similarly from the right. This shows that $\sigma_{\mathcal{A}}(a) \setminus \{0\} \subset \sigma_{\mathcal{B}}(a)$. Now consider the C*-subalgebra $\mathcal{B} + \mathbb{C}(I_{\mathcal{A}} - I_{\mathcal{B}})$ of $\mathcal{A}$. It is unital with the same unit as $\mathcal{A}$, so by the first part of the proof, $\sigma_{\mathcal{B}+\mathbb{C}(I_{\mathcal{A}}-I_{\mathcal{B}})}(a) = \sigma_{\mathcal{A}}(a)$. If $\lambda \neq 0$ and $a - \lambda I_{\mathcal{A}}$ is invertible in $\mathcal{B} + \mathbb{C}(I_{\mathcal{A}} - I_{\mathcal{B}})$, there exists $c + \mu(I_{\mathcal{A}} - I_{\mathcal{B}}) \in \mathcal{B} + \mathbb{C}(I_{\mathcal{A}} - I_{\mathcal{B}})$ such that

$$I_{\mathcal{A}} = \big(c + \mu(I_{\mathcal{A}} - I_{\mathcal{B}})\big)(a - \lambda I_{\mathcal{A}}) = c(a - \lambda I_{\mathcal{B}}) + \mu\lambda(I_{\mathcal{A}} - I_{\mathcal{B}}).$$

Multiplying by $I_{\mathcal{A}} - I_{\mathcal{B}}$ we get $\mu\lambda = 1$, and then $c(a - \lambda I_{\mathcal{B}}) = I_{\mathcal{B}}$. With a similar computation on the right, we conclude that $\sigma_{\mathcal{B}}(a) \setminus \{0\} \subset \sigma_{\mathcal{A}}(a)$. Thus $\sigma_{\mathcal{A}}(a) \setminus \{0\} = \sigma_{\mathcal{B}}(a) \setminus \{0\}$, and we it follows that $\sigma_{\mathcal{A}}(a) = \sigma_{\mathcal{B}}(a) \cup \{0\}$.

Now suppose that $\mathcal{B}$ is not unital. The spectrum $\sigma_{\mathcal{B}}(a)$ is defined in the unitization $\tilde{\mathcal{B}}$. We may replace $\mathcal{A}$ with its unitization if $\mathcal{A}$ is not unital, and then unitize $\mathcal{B}$ with the unit $I_{\mathcal{A}}$. Therefore both $\sigma_{\mathcal{A}}(a)$ and $\sigma_{\mathcal{B}}(a)$ are defined in an algebra and subalgebra with the same unit; and by the first part of the proof we have $\sigma_{\mathcal{B}}(a) = \sigma_{\mathcal{A}}(a)$. $\square$

***11.1.13. Proposition.***

*Let $a \in \mathcal{A}$. Then there exist unique $b, c \in \mathcal{A}^{\mathrm{sa}}$ such that $a = b + ic$.*

*Proof.* Given $a \in \mathcal{A}$, let

$$b = \frac{a + a^*}{2}, \qquad c = \frac{a - a^*}{2i}. \tag{11.5}$$

It is straightforward to check that $b, c \in \mathcal{A}^{\mathrm{sa}}$ and that $a = b + ic$. If $b_1 + ic_1 = b_2 + ic_2$ with $b_1, b_2, c_1, c_2$ selfadjoint, then $b_1 - b_2 = i(c_1 - c_2)$. As the left-hand-side is selfadjoint, taking adjoints we get $i(c_1 - c_2) = -i(c_1 - c_2)$, so $c_1 = c_2$, and this forces $b_1 = b_2$. $\square$

**11.1.14. Definition.**

*The elements $b, c$ constructed in (11.5) are called the* ***real*** *and* ***imaginary*** *parts of $a$, often denoted by $\operatorname{Re} a$ and $\operatorname{Im} a$.*

**11.1.1. Exercises.**

(11.1.1) Prove the statements made in Remark 11.1.3.

(11.1.2) Show that the norm defined in (11.1) is submultiplicative.

(11.1.3) Show that if $x \in \tilde{\mathcal{A}}$ satisfies $bx = 0$ for all $b \in \mathcal{A}$, then $x = 0$.

(11.1.4) Show that if $\mathcal{A}$ is a C*-algebra, unital or not, then

$$\|(a, \lambda)\| = \max\left\{|\lambda|, \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\}\right\}.$$

defines a norm that makes $\tilde{\mathcal{A}}$ a C*-algebra, and that when $\mathcal{A}$ is not unital we recover $\tilde{\mathcal{A}}$ as in Proposition 11.1.4. *(Hint: at some point you will possibly need the fact that the norm on a C*-algebra is unique)*

(11.1.5) Show that the inequality

$$|\lambda| \leq \sup\{\|ab + \lambda b\| : \ \|b\| \leq 1\}, \qquad a \in \mathcal{A},$$

holds when $\mathcal{A}$ is non-unital, and it fails when $\mathcal{A}$ is unital.

(11.1.6) Show that, for any $a \in \mathcal{A}$, $\sigma(a^*) = \overline{\sigma(a)}$.

(11.1.7) Let $\{\mathcal{A}_j\}_{j \in J}$ be a family of C*-subalgebras of a $C^*$-algebra $\mathcal{A}$. Show that $\bigcap_j \mathcal{A}_j$ is a C*-subalgebra of $\mathcal{A}$.

(11.1.8) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Show that $C^*(a)$ is equal to

$$\overline{\{p(a, a^*) : \ p \ \text{ non-commutative polynomial with } p(0,0) = 0\}}.$$

(11.1.9) Let $\mathcal{A}$ be a C*-algebra such that every normal element has spectrum consisting of a single point. Prove that $\mathcal{A} = \mathbb{C}$.

(11.1.10) Show that if $\mathcal{A}$ is not unital then $(\tilde{\mathcal{A}})^* \simeq \mathcal{A}^* \oplus_1 \mathbb{C}$.

## 11.2. Continuous Functional Calculus

The focus on this section is on abelian C*-algebras. These can be characterized thoroughly. For that, we need to pay attention to the characters $\Sigma(\mathcal{A})$ of an abelian C*-algebra $\mathcal{A}$ (Definition 9.2.22). From Proposition 9.2.23, there is a one-to-one correspondence between the characters of $\mathcal{A}$ and its maximal ideals (we only consider bilateral closed ideals). We remark that characters of the abelian C*-algebra $C_0(T)$ played a crucial role in Section 7.4.

Being a subset of $\mathcal{A}^*$, we will endow $\Sigma(\mathcal{A})$ with the weak*-topology; this makes it a locally compact Hausdorff topological space.

It is worth mentioning that in general characters are rare: for instance $M_n(\mathbb{C})$, for $n \geq 2$, has no characters (Exercise 9.2.13). But they are abundant in abelian C*-algebras: enough to separate points, in fact. We can get this from Proposition 9.2.24. Indeed, given $a \in \mathcal{A}$ nonzero, then $\|a\|^2 = \|a^*a\| > 0$. We have that $\|a^*a\| = \operatorname{spr}(a^*a)$ (Lemma 11.1.10), and so by Proposition 9.2.24 there exists $\tau \in \Sigma(\mathcal{A})$ with $|\tau(a^*)\tau(a)| = |\tau(a^*a)| = \|a^*a\|$; in particular, $\tau(a) \neq 0$.

***11.2.1. Proposition.***

*Let $\mathcal{A}$ be abelian, $\tau \in \Sigma(\mathcal{A})$, and $a \in \mathcal{A}$. Then*

*(i)* $\|\tau\| = 1$;

*(ii) If* $a^* = a$, *then* $\tau(a) \in \mathbb{R}$;

*(iii)* $\tau(a^*) = \overline{\tau(a)}$.

*Proof.*

(i) Suppose that there exists $a \in \mathcal{A}$ with $\|a\| < 1$ and $\tau(a) = 1$. Let $b = \sum_{n=1}^\infty a^n \in \mathcal{A}$. As $a + ab = b$,

$$\tau(b) = \tau(a) + \tau(a)\tau(b) = 1 + \tau(b),$$

a contradiction. So $\|\tau\| \leq 1$. Also, if we fix $a \in \mathcal{A}$ with $\|a\| = 1$, as just mentioned there exists $\tau \in \Sigma(\mathcal{A})$ with $|\tau(a^*a)| = 1$. Then

$$1 = |\tau(a^*a)| \leq \|\tau\| \, \|a^*a\| = \|\tau\| \, \|a\|^2 = \|\tau\|.$$

Hence $\|\tau\| = 1$.

(ii) We can extend $\tau$ to the unitization $\tilde{\mathcal{A}}$ by $\tilde{\tau}(a,\lambda) = \tau(a) + \lambda$; this is still nonzero, linear, and multiplicative. This extension is unique, because a character in a unital algebra has to satisfy $\tau(I_{\tilde{\mathcal{A}}}) = 1$. So we may assume that $\mathcal{A}$ is unital. We assume that $a = a^*$. Let $v = e^{ia}$. We can define this via the series for the exponential, but in any case it will agree with the holomorphic functional calculus as in Theorem 9.3.2. From $a = a^*$ and using the series representation for $e^{ia}$ we get that $v^* = e^{-ia}$ (Exercise 11.2.1). Since $1 = e^{iz}e^{-iz}$, the holomorphic functional calculus gives us $v^*v = vv^* = I_{\mathcal{A}}$. So $v$ is a unitary; combining Propositions 11.1.7 and 9.2.9 we have that $\lambda \in \sigma(v)$ if and only if $\lambda^{-1} \in \sigma(v^*)$; since $\|v\| = \|v^*v\|^{1/2} = 1$ these imply respectively $|\lambda| \leq 1$ and $|\lambda| \geq 1$, and so $\sigma(v) \subset \mathbb{T}$. Combining this with the Spectral Mapping Theorem (Theorem 9.3.3), $e^{i\tau(a)} \in \mathbb{T}$, which in turn implies that $\tau(a) \in \mathbb{R}$.

(iii) We have $a = b + ic$ with $b^* = b$, $c^* = c$, where $b = (a + a^*)/2$ and $c = -i(a - a^*)/2$. Then, as $\tau(b), \tau(c) \in \mathbb{R}$,

$$\tau(a^*) = \tau(b - ic) = \tau(b) - i\tau(c) = \overline{\tau(b) + i\tau(c)} = \overline{\tau(a)}. \qquad \square$$

When $\mathcal{A} = C_0(T)$, with $T$ locally compact Hausdorff, $\Sigma(\mathcal{A})$ is easy to find. Indeed, we proved in Proposition 7.4.7 that $\Sigma(\mathcal{A}) = \{\tau_t :\ t \in T\}$, where $\tau_t(f) = f(t)$. Two other ways of proving so are mentioned in the paragraph before Proposition 7.4.13.

***11.2.2. Proposition.***

*If $\mathcal{A}$ is unital, then $\Sigma(\mathcal{A})$ is compact.*

*Proof.* Let $\{\tau_j\}$ be a net in $\Sigma(\mathcal{A})$. By Proposition 11.2.1, the net is contained in the unit ball of $\mathcal{A}^*$; by Banach–Alaoglu (Theorem 7.2.13), there exists subnet converging to $\tau \in \overline{B_1^{\mathcal{A}^*}(0)}$. As the limit is pointwise, it is apparent that $\tau$ is linear and multiplicative. From $\mathcal{A}$ unital, we have that $\tau_j(I_{\mathcal{A}}) = I_{\mathcal{A}}$ for all $j$, so $\tau(I_{\mathcal{A}}) = I_{\mathcal{A}}$ and $\tau \neq 0$, so $\tau \in \Sigma(\mathcal{A})$. $\square$

When $\mathcal{A}$ is abelian and non-unital, we will soon see that $\Sigma(\mathcal{A})$ is not compact. It follows from Proposition 11.2.2 that a non-zero limit of a net of characters is always a character; hence if $\Sigma(\mathcal{A})$ is not compact, then $\overline{\Sigma(\mathcal{A})} = \Sigma(\mathcal{A}) \cup \{0\}$. For an example of a net—actually, a sequence in this case—of characters converging to zero, let $\mathcal{A} = c_0$, and $\tau_n(a) = a(n)$. Then $\tau_n \to 0$ in the weak*-topology.

The Gelfand Transform (Theorem 11.2.3) gives a complete characterization of abelian C*-algebras. It has far-reaching consequences, as we will soon see. To follow the proof below, the canonical example is $\mathcal{A} = C_0(\mathbb{R})$, where $\Sigma(\mathcal{A}) = \mathbb{R}$ as point evaluations; and $0 \in \overline{\Sigma(\mathcal{A})}$ because $\tau_n \to 0$, where $\tau_n(f) = f(n)$.

***11.2.3. Theorem. (Gelfand Transform)***

*Let $\mathcal{A}$ be an abelian $C^*$-algebra. Then $\mathcal{A} \simeq C_0(\Sigma(\mathcal{A}))$.*

*Proof.* If $\Sigma(\mathcal{A})$ is compact, then $C_0(\Sigma(\mathcal{A})) = C(\Sigma(\mathcal{A}))$. Otherwise, as mentioned above we have $\overline{\Sigma(\mathcal{A})} = \Sigma(\mathcal{A}) \cup \{0\}$. The set $K = \{\tau : \ |f(\tau)| \geq \varepsilon\}$ is weak*-closed by the continuity of $f$, and hence compact. So the continuous functions on $\Sigma(\mathcal{A})$ that vanish at infinity are precisely those such that $f(\tau_j) \to 0$ when $\tau_j \to 0$.

Define $\Gamma : \mathcal{A} \to C_0(\Sigma(\mathcal{A}))$ by $\Gamma(a) = \hat{a}$, where $\hat{a}(\tau) = \tau(a)$. This $\hat{a}$ is trivially continuous: if $\tau_j \to \tau$, then $\hat{a}(\tau_j) = \tau_j(a) \to \tau(a) = \hat{a}(\tau)$. If $\tau_j \to 0$, then $\hat{a}(\tau_j) = \tau_j(a) \to 0$, so $\hat{a} \in C_0(\Sigma(\mathcal{A}))$. The fact that $\tau$ is linear, multiplicative, and $*$-preserving guarantees that $\Gamma$ is a $*$-homomorphism.

Next we see that $\Gamma$ is isometric. Given $a \in \mathcal{A}$ we have that $a^*a$ is selfadjoint. By Lemma 11.1.10, $\|a^*a\| = \operatorname{spr}(a^*a)$; and by Proposition 9.2.24, there exists $\tau \in \Sigma(\mathcal{A})$ with $|\tau(a^*a)| = \|a^*a\|$. Then, using that $\tau$ preserves adjoints (Proposition 11.2.1),

$$|\hat{a}(\tau)|^2 = |\tau(a)|^2 = \overline{\tau(a)}\tau(a) = \tau(a^*a) = \|a^*a\| = \|a\|^2,$$

so $|\hat{a}(\tau)| = \|a\|$. This shows that $\|\hat{a}\| = \|a\|$, and hence $\Gamma$ is isometric; in particular, it is injective. Being isometric, the image of $\Gamma$ is closed: indeed, if $\{\hat{a}_j\}$ is Cauchy, then $\|a_j - a_k\| = \|\hat{a}_j - \hat{a}_j\|$ and so $\{a_j\}$ is Cauchy; letting $a = \lim a_j$ we have $\hat{a} = \lim \hat{a}_j$.

Finally, the set $\Gamma(\mathcal{A}) = \{\Gamma(a) : \ a \in \mathcal{A}\} \subset C_0(\Sigma(\mathcal{A}))$ is a closed, selfadjoint subalgebra that separates points and vanishes nowhere (Exercise 11.2.2). By Stone–Weierstrass (Corollary 7.4.23), $\Gamma(\mathcal{A}) = C_0(\Sigma(\mathcal{A}))$. □

**11.2.4. Definition.**

*The **Gelfand Transform** is the map $\Gamma : \mathcal{A} \to C_0(\Sigma(\mathcal{A}))$ from the proof of Theorem 11.2.3. The name is sometimes also applied to the related map in the proof of Corollary 11.2.6 below.*

*11.2.5. Remark.* Here is an interesting thing that happens due to the availability of the Gelfand transform. Consider $\mathcal{A} = L^\infty[0,1]$ (we could have used any $\sigma$-finite measure space $(X, \mu)$, but let us go with a familiar one). Then $\mathcal{A}$ is an abelian C*-algebra. The algebra $\mathcal{A}$ contains many non-continuous functions like characteristic functions and even much worse-behaved functions, which are not the kind of thing we think about when we say "continuous". But Theorem 11.2.3

says that $\mathcal{A}$ (being unital) is C*-isomorphic to $C(K)$ for a certain compact Hausdorff space $K$. So, on one side of the isomorphism $L^\infty[0,1] \simeq C(K)$ we have functions that do not look continuous at all, and on the other side all functions are continuous! What kind of weird space is this mysterious $K$? We already know, actually. We showed via Proposition 9.9.7 and Theorem 9.9.14 that $\mathcal{A}$ is isometrically isomorphic, as Banach spaces, to $C(M)$ where $M$ is an extremally disconnected space. So $C(K)$ and $C(M)$ are isometrically isomorphic as Banach spaces; by Banach–Stone (Theorem 9.4.11) we have that $K$ and $M$ are homeomorphic. It is worth mentioning that it is possible to show that $K$ is extremally disconnected without appealing to Chapter 9, by focusing on order properties of $\mathcal{A}$; this can be argued both using measure theory and using von Neumann algebras.

Next we establish the continuous functional calculus for normal elements in unital C*-algebra. Taking $\mathcal{A}$ to be unital is not a big deal in this case, as we can work on the unitization $\tilde{\mathcal{A}}$ when $\mathcal{A}$ is non-unital. There is a possible source of confusion in $C(\sigma(a))$ below between the **identity function** $z \longmapsto z$ and the identity of the algebra, which is the constant function 1.

***11.2.6. Corollary. (Continuous Functional Calculus)***

*Let $\mathcal{A}$ be a unital C*-algebra and $a \in \mathcal{A}$ be normal. Then $C(\sigma(a)) \simeq C^*(I_{\mathcal{A}}, a)$, via a $*$-isomorphism that maps the identity function $z \longmapsto z$ to $a$.*

*Proof.* The normality of $a$ is what makes $C^*(I_{\mathcal{A}}, a)$ abelian. We know from Theorem 11.2.3 that $C^*(I_{\mathcal{A}}, a) \simeq C(T)$, where $T = \Sigma(C^*(I_{\mathcal{A}}, a))$ is compact. We will prove that $T$ is homeomorphic to $\sigma(a)$. Let $\gamma : T \to \sigma(a)$ be the map $\gamma(\tau) = \tau(a)$. This map is well defined and onto by Proposition 9.2.24. It is continuous: by definition of the weak*-topology, if $\tau_j \to \tau$ then $\gamma(\tau_j) = \tau_j(a) \to \tau(a) = \gamma(\tau)$. It is injective: if $\tau(a) = \eta(a)$ for $\tau, \eta \in T$, then $\tau(p(a,a^*)) = \eta(p(a,a^*))$ for all two-variable polynomials $p$; so they agree on a dense subset of $C^*(I_{\mathcal{A}}, a)$, and by continuity we get $\tau = \eta$. It remains to check that $\gamma^{-1}$ is continuous. If $\tau_j(a) \to \tau(a)$, then $\tau_j(p(a,a^*)) \to \tau(p(a,a^*))$ for all two-variable polynomials $p$. If $b \in C^*(I_{\mathcal{A}}, a)$ and $\varepsilon > 0$, there exists $p$ with $\|b - p(a,a^*)\| < \varepsilon$. Then

$$\begin{aligned}
|\tau_j(b) - \tau(b)| &\le |\tau_j(b) - \tau_j(p(a,a^*))| + |\tau_j(p(a,a^*)) - \tau(p(a,a^*))| \\
&\qquad + |\tau(p(a,a^*)) - \tau(b)| \\
&\le 2\|p(a,a^*) - b\| + |\tau_j(p(a,a^*)) - \tau(p(a,a^*))| \\
&\le 2\varepsilon + |\tau_j(p(a,a^*)) - \tau(p(a,a^*))|.
\end{aligned}$$

Then $\limsup_j |\tau_j(b) - \tau(b)| \leq 2\varepsilon$ and, as $\varepsilon$ was arbitrary, $\tau_j(b) \to \tau(b)$.

As in Theorem 7.4.8, the homeomorphism $\gamma$ induces a $*$-isomorphism $\tilde{\gamma} : C(\sigma(a)) \to C(T)$ by $(\tilde{\gamma} f)(\tau) = f(\gamma(\tau))$.

If we now look at the $*$-isomorphism $\Gamma^{-1} \circ \tilde{\gamma} : C(\sigma(a)) \to C^*(a)$, we have that

$$\tilde{\gamma}(\mathrm{id}_{C(\sigma(a))})(\tau) = \mathrm{id}_{C(\sigma(a))}(\gamma(\tau)) = \gamma(\tau) = \tau(a) = \hat{a}(\tau).$$

Then $\tilde{\gamma}(\mathrm{id}_{C(\sigma(a))}) = \hat{a}$ and so

$$(\Gamma^{-1} \circ \tilde{\gamma})(\mathrm{id}_{C(\sigma(a))}) = \Gamma^{-1}\big(\tilde{\gamma})(\mathrm{id}_{C(\sigma(a))})\big) = \Gamma^{-1}(\hat{a}) = a. \qquad \square$$

The way we use Corollary 11.2.6 is to note that since the isomorphism maps the identity function to $a$, it maps polynomials $p$ to $p(a)$. Then we *define*:

**11.2.7. Definition. (Continuous Functional Calculus)**

*Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$ be normal. Then, if $\Theta : C(\sigma(a)) \to C^*(a)$ denotes the $*$-isomorphism from Corollary 11.2.6, we define*

$$f(a) = \Theta(f).$$

As $\Theta(p) = p(a)$ for all complex polynomials and $\Theta$ is continuous, if $f = \lim_n p_n$ uniformly then $f(a) = \lim_n p(a)$. Even more generally, if $f_n \to f$ uniformly on $C(\sigma(a))$, then $f_n(a) \to f(a)$. Also, $\Theta$ is a $*$-homomorphism, so we have things like $(f+g)(a) = f(a) + g(a)$, $(fg)(a) = f(a)g(a)$. This is used constantly in C*-algebra manipulations.

As opposed to arbitrary Banach algebras, we showed in Proposition 11.1.12 that the spectrum of an element in a C*-algebra does not change if we pass to a C*-subalgebra. This fact is crucial in developing and using functional calculus in C*-algebras, as any construction that depends on the spectrum of $a$ can be carried on $C^*(a)$ and works on any C*-algebra that contains $a$.

***11.2.8. Corollary. (Spectral Mapping: Continuous Version)***

*Let $a \in \mathcal{A}$ be normal, $f \in C(\sigma(a))$. Then*

$$\sigma(f(a)) = f(\sigma(a)).$$

*Proof.* Since $C^*(a) \simeq C(\sigma(a))$ by Corollary 11.2.6,

$$\sigma(f(a)) = \sigma(f) = f(\sigma(a)).$$

The second equality is Exercise 9.2.5. $\square$

*11.2.9. Example.* When $\mathcal{A} = \ell^\infty(\mathbb{N})$, by Theorem 11.2.3 we have $\mathcal{A} \simeq C(X)$, where $X$ is the spectrum of $\mathcal{A}$. Since $\mathcal{A}$ is unital, $X$ is compact. We will show that $X = \beta\mathbb{N}$, the Stone-Čech compactification of $\mathbb{N}$.

By Theorem 7.4.14 each $x \in \ell^\infty(\mathbb{N}) = C_b(\mathbb{N})$ admits a unique extension $\tilde{x} \in C(\beta\mathbb{N})$. Hence the map $x \longmapsto \tilde{x}$ from $\ell^\infty(\mathbb{N})$ onto $C(\beta\mathbb{N})$ is a $*$-isomorphism. By Theorem 7.4.8, $X = \Sigma(\ell^\infty(\mathbb{N}))$ is homeomorphic to $\beta\mathbb{N}$.

We already knew from Proposition 9.9.2 and Theorem 9.9.14 that $\mathcal{A}$ is isometrically isomorphic, as a Banach space, to $C(M)$ with $M$ extremally disconnected. By Banach–Stone (Theorem 9.4.11) we have that $M$ is homeomorphic to $\beta\mathbb{N}$.

Here is a second step in understanding $*$-homomorphisms, and our first use of the continuous functional calculus.

**11.2.10. Proposition.**

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then $\pi$ is bounded, and $\|\pi\| \leq 1$. If $\pi$ is injective, then it is isometric.*

*Proof.* The first part of the result was established in Proposition 11.1.11.

If $\pi$ is injective, and $\|\pi(a)\| < \|a\|$ for some $a$, we have $\|\pi(a^*a)\| < \|a^*a\|$. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous with $f|_{[0,\|\pi(a^*a)\|]} = 0$, $f(\|a^*a\|) = 1$. From Corollary 11.2.8, $\sigma(f(a^*a)) = f(\sigma(a^*a)) \ni f(\|a^*a\|) = 1$, so $f(a^*a) \neq 0$. But since $f|_{\sigma(\pi(a^*a))} = 0$, we have $0 = f(\pi(a^*a)) = \pi(f(a^*a))$, a contradiction since $\pi$ is injective. It follows that $\|\pi(a)\| = \|a\|$ for all $a \in \mathcal{A}$. □

**11.2.11. Corollary.**

*Let $\mathcal{A}$, $\mathcal{B}$ be abelian C*-algebras. The following statements are equivalent:*

*(i) $\mathcal{A}$ and $\mathcal{B}$ are isometrically isomorphic as Banach spaces;*

*(ii) $\mathcal{A}$ and $\mathcal{B}$ are isomorphic as C*-algebras.*

*Proof.* Because of Proposition 11.2.10 we only need to show (i) $\Longrightarrow$ (ii). By Theorem 11.2.3 we can write $\mathcal{A} = C_0(S)$, $\mathcal{B} = C_0(T)$. If $C_0(S)$ and $C_0(T)$ are isometrically isomorphic as Banach spaces, then by Banach–Stone (Theorem 9.4.11) we get that $S$ and $T$ are homeomorphic. Then $C_0(S)$ and $C_0(T)$ are isomorphic as C*-algebras by Gelfand–Kolmogorov (Theorem 7.4.8). □

We can now fulfill a promise made long ago in Remark 9.9.17. It is obvious that $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$ are not isomorphic as C*-algebras, as the former has minimal projections and the latter does not. But below we are talking of isomorphism as Banach spaces!

**11.2.12. Corollary.**

*There is no isometric isomorphism between the Banach spaces $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$.*

*Proof.* By Corollary 11.2.11, it is enough to show that $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$ are not isomorphic as C*-algebras. Consider $e_1 \in \ell^\infty(\mathbb{N})$. We have that $e_1$ is minimal, in the sense that $e_1\,\ell^\infty(\mathbb{N}) = \mathbb{C}\,e_1$. If $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$ were isomorphic as C*-algebras, there would exist a projection $p \in L^\infty[0,1]$ with the same property. Projections in a function space are necessarily characteristic functions. So there would have exist a measurable set $E$ of positive Lebesgue measure with $1_E\,L^\infty[0,1] = \mathbb{C}\,1_E$. But using Exercise 2.3.17 we can write $E = E_1 \cup E_2$, both measurable with positive Lebesgue measure, and so $1_{E_1} \in 1_E\,L^\infty[0,1]$ but it is not a scalar multiple of $1_E$. ☐

It is possible to do a version of Corollary 11.2.6 in the non-unital case. Namely, with the notation $C_{\text{zero}}(T) = \{f \in C(T) :\ f(0) = 0\}$,

**11.2.13. Corollary. (Continuous Functional Calculus)**

*Let $\mathcal{A}$ be a non-unital C*-algebra and $a \in \mathcal{A}$ be normal. Then $C^*(a) \simeq C_{\text{zero}}(\sigma(a))$, via a $*$-isomorphism that maps $a$ to the identity function $z \longmapsto z$.*

*Proof.* Applying Corollary 11.2.6 in the unitization of $C^*(a)$ we obtain an isomorphism $C^*(I_{\mathcal{A}}, a) \simeq C(\sigma(a))$ that maps $a$ to the identity $z \longmapsto z$. The restriction of this isomorphism to $C^*(a)$ has as its image the C*-algebra generated by the identity (note that the range of the restriction of the isomorphism to $C^*(a)$ is a C*-algebra since it is an isometry). The C*-algebra generated by the identity inside $C(\sigma(a))$ is $C_{\text{zero}}(\sigma(a))$. This can be seen by noticing that the algebra generated by the identity is the algebra of non-commutative polynomials $p$ with $p(0,0) = 0$. Together with their adjoints we get a dense $*$-algebra by Stone–Weierstrass (Theorem 7.4.20). ☐

What we get from Corollary 11.2.13 is that when $\mathcal{A}$ is not unital and $a \in \mathcal{A}$ is normal, we can do continuous functional calculus (and everything we said so far about it applies) with the exception that we can only consider $f(a)$ for $f \in C(\sigma(a))$ with $f(0) = 0$.

The isometry of the functional calculus also gives us that $\|f(a)\| = \|f\|_\infty$. In particular, for $a$ normal we recover the result from Lemma 11.1.10,

$$\|a\| = \|z\|_\infty = \operatorname{spr}(a), \tag{11.6}$$

where $z(t) = t$ is the identity function on $\sigma(a)$.

**11.2.1. Exercises.**

(11.2.1) Let $a \in \mathcal{A}$ be selfadjoint and let $v = e^{ia}$. Show that $v^* = e^{-ia}$.

(11.2.2) Show that if $\mathcal{A}$ is an abelian C$^*$-algebra then $\Gamma(A) = \{\Gamma(a) : a \in A\} \subset C_0(\Sigma(\mathcal{A}))$ is a closed, selfadjoint subalgebra that separates points and vanishes nowhere.

(11.2.3) Let $\mathcal{A}$ be a C$^*$-algebra, $a \in \mathcal{A}$ and $f \in C(\sigma(a^*a) \cup \{0\})$. Show that $af(a^*a) = f(aa^*)a$.

(11.2.4) Let $\mathcal{A}$ be a C$^*$-algebra and $a \in \mathcal{A}$ be normal. Let $\lambda \in \mathbb{C} \setminus \sigma(a)$. Show that $\operatorname{dist}(\lambda, \sigma(a)) = \|(a - \lambda I_{\mathcal{A}})^{-1}\|^{-1}$.

(11.2.5) Let $\mathcal{A}, \mathcal{B}$ be C$^*$-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Let $\hat{\mathcal{A}}$ denote $\mathcal{A}$ is $\mathcal{A}$ is unital, and $\tilde{\mathcal{A}}$ if $\mathcal{A}$ is not unital, and do similarly with $\mathcal{B}$. Show that there exists a unique $*$-homomorphism $\tilde{\rho} : \hat{\mathcal{A}} \to \hat{\mathcal{B}}$, unital onto its image, that extends $\rho$.

(11.2.6) Let $\mathcal{A}, \mathcal{B}$ be C$^*$-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Prove that for all $a \in \mathcal{A}$, $\sigma(\rho(a)) \subset \sigma(a) \cup \{0\}$. The zero can be omitted when $\rho(\mathcal{A})$ is not unital, and also when $\mathcal{A}, \mathcal{B}$ are unital and $\rho$ is unital. The zero cannot be omitted in general.

(11.2.7) Let $\mathcal{A}, \mathcal{B}$ be C$^*$-algebras and $\rho : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Given $a \in \mathcal{A}$ normal and $f \in C(\sigma(a) \cup \{0\})$, show that $f(\rho(a))$ makes sense and that $f(\rho(a)) = \rho(f(a))$.

(11.2.8) The Banach algebra $\mathbb{A} = \{f \in C(\overline{\mathbb{D}}) : \text{ analytic on } \mathbb{D}\}$ is not a C$^*$-subalgebra of $C(\overline{\mathbb{D}})$ because it does not contain adjoints for non-scalar functions. Show that $\mathbb{A}$ fails to be a C$^*$-algebra in a deeper sense: it is not isomorphic to $C(X)$ for any compact Hausdorff $X$, even if the isomorphism is not required to be isometric.

(11.2.9) Let $v \in \mathcal{A}$ be a selfadjoint partial isometry (that is, $v^* = v$, and $v^2$ is a projection). Show that there exist projections $p, q \in \mathcal{A}$ with $pq = 0$ and $v = p - q$.

## 11.3. Positivity

We begin with three basic definitions, modelled on the corresponding notions for operators. We have already mentioned normal operators in Definition 11.1.9.

**11.3.1. Definition.**

*Given $a \in \mathcal{A}$, we say that $a$ is*

*(i)* ***normal*** *if $a^*a = aa^*$;*

*(ii)* ***selfadjoint****, if $a^* = a$;*

*(iii)* ***positive****, if $a = b^*b$ for some $b \in \mathcal{A}$.*

*We will denote the sets of selfadjoint and positive elements of $\mathcal{A}$ by $\mathcal{A}^{\mathrm{sa}}$ and $\mathcal{A}^+$, respectively.*

*For selfadjoint $a, b \in \mathcal{A}$, we say that $a \leq b$ if $b - a \in \mathcal{A}^+$.*

These notions were already considered in the case where $\mathcal{A} = \mathcal{B}(\mathcal{H})$ in Section 10.4. And of course the new notions agree with the old ones, though positivity is defined differently here. But an arbitrary C*-algebra does not necessarily consist of operators on a Hilbert space—for now!—so we need to treat these notions on their own.

The order just defined is intrinsically related to the positive cone $\mathcal{A}^+$; for $a \geq 0$ if and only if $a \in \mathcal{A}^+$. To see a familiar case where this arises, consider the situation where $\mathcal{A} = \mathbb{C}$; then $\mathcal{A}^{\mathrm{sa}} = \mathbb{R}$ and $\mathcal{A}^+ = [0, \infty)$, and $a \leq b$ is the usual order of numbers. When $\mathcal{A} = M_n(\mathbb{C})$ we have that $\mathcal{A}^{\mathrm{sa}}$ are the conjugate-symmetric matrices ($a_{jk} = \overline{a_{kj}}$), and $\mathcal{A}^+$ are the positive-semidefinite matrices.

***11.3.2. Lemma.***

*Let $a \in \mathcal{A}$. The following statements are equivalent:*

*(i) $a \geq 0$;*

*(ii) $a$ is normal and $\sigma(a) \subset [0, \infty)$.*

*Proof.* (i) $\implies$ (ii) By definition, $a = b^*b$ for some $b \in \mathcal{A}$, so $a = a^*$. Let $\mathcal{A}_0 \subset \mathcal{A}$ be the C*-subalgebra (unital if $\mathcal{A}$ is unital) generated by $a$. By Proposition 11.1.12, $\sigma(a)$ is the same in either algebra. For any $\tau \in \Sigma(\mathcal{A}_0)$ we have $\tau(a) = \tau(b^*b) = |\tau(b)|^2 \geq 0$ by Proposition 11.2.1. So $\sigma(a) \subset [0, \infty)$ by Proposition 9.2.24.

(ii) $\implies$ (i) Since $a$ is normal and $\sigma(a) \subset [0, \infty)$, the function $f(t) = t^{1/2}$ is defined and continuous on $\sigma(a)$. Let $b = f(a)$. Then, since the continuous functional calculus is a homomorphism, $b^2 = (f(a))^2 = f^2(a) = a$, so $a = b^2 = b^*b \geq 0$. $\square$

*11.3.3. Remark* . In the case where $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$, Lemma 11.3.2 was done in Proposition 10.4.7. Thing is, here we are studying C$^*$-algebras without assuming that they live inside some $\mathcal{B}(\mathcal{H})$. With enough theory developed, it will be possible to show that any C$^*$-algebra can be represented inside some $\mathcal{B}(\mathcal{H})$ (Corollary 11.6.6).

*11.3.4. Remark.* It is worth noticing that the characterizations in Lemma 11.3.2 are not equivalent for a $*$-subalgebra of a C$^*$-algebra. That is, if $\mathcal{A}_0 \subset \mathcal{A}$ is a $*$-subalgebra of the C$^*$-algebra $\mathcal{A}$, the sets $\{a \in \mathcal{A}_0 : \ a \geq 0\}$ and $\{b^*b : \ b \in \mathcal{A}_0\}$ are not necessarily equal. Consider $\mathcal{A} = C[0,1]$ and $\mathcal{A}_0 = \mathbb{C}[x]$. Then the polynomial $a$ given by $a(t) = t$ is positive in $\mathcal{A}$ (for $a^* = a$ and $\sigma(a) = [0,1]$) but there is no $b \in \mathcal{A}_0$ with $b^*b = a$ (since the square root function is not a polynomial).

The $b$ obtained in the proof of Lemma 11.3.2 is unique (Exercise 11.3.3) and it will be denoted by $a^{1/2}$. We refer to Remark 10.4.5 on how not practical it is to find square roots explicitly in the case of operators. But, as already seen in Chapter 10, their use for theoretical results is widespread and crucial.

The next few results will be used often and without mention, as they are the bread and butter of C$^*$-algebraic manipulations.

***11.3.5. Proposition.***

*Let $a, b \in \mathcal{A}^{\mathrm{sa}}$, $c \in \mathcal{A}$. If $a \leq b$, then $cac^* \leq cbc^*$.*

*Proof.* Since $b-a \geq 0$, there exists $d \in \mathcal{A}$ with $b-a = d^*d$. Then $c(b-a)c^* = cd^*dc^* = (dc^*)^*dc^* \geq 0$. ☐

***11.3.6. Corollary.***

*Let $a \in \mathcal{A}$ be selfadjoint, $\alpha, \beta \in \mathbb{R}$. The following statements are equivalent:*

*(i)* $\alpha\, I_{\mathcal{A}} \leq a \leq \beta\, I_{\mathcal{A}}$;

*(ii)* $\sigma(a) \subset [\alpha, \beta]$.

*Proof.* (i) $\implies$ (ii) Since $a - \alpha\, I_{\mathcal{A}} \geq 0$, by Lemma 11.3.2 we have $\sigma(a) - \alpha = \sigma(a - \alpha\, I_{\mathcal{A}}) \subset [0,\infty)$, so $\sigma(a) \subset [\alpha, \infty)$. Similarly, from $\beta\, I_{\mathcal{A}} - a \geq 0$ we get $\beta - \sigma(a) = \sigma(\beta\, I_{\mathcal{A}} - a) \subset [0,\infty)$, so $\sigma(a) \subset (-\infty, \beta]$. Hence $\sigma(a) \subset [\alpha, \beta]$.

(ii) $\implies$ (i) We can basically reverse the steps above. We have $\sigma(a - \alpha I) = \sigma(a) - \alpha \subset [0, \beta - \alpha]$, so by Lemma 11.3.2 we get $\alpha I_{\mathcal{A}} \leq a$. With a similar argument we get $a \leq \beta I_{\mathcal{A}}$. $\square$

**11.3.7. Corollary.**

*Let $\mathcal{A}$ be unital and $a, b \in GL(\mathcal{A})$. If $a \leq b$, then $b^{-1} \leq a^{-1}$.*

*Proof.* By Proposition 11.3.5, we have $0 \leq b^{-1/2}ab^{-1/2} \leq I_{\mathcal{A}}$. Then

$$\sigma(b^{-1/2}ab^{-1/2}) \subset [0,1]$$

by Corollary 11.3.6. By Proposition 9.2.15, since

$$b^{-1/2}ab^{-1/2} = (a^{1/2}b^{-1/2})^* a^{1/2}b^{-1/2},$$

$\sigma(a^{1/2}b^{-1}a^{1/2}) \subset [0,1]$. This implies, via Corollary 11.3.6, that

$$0 \leq a^{1/2}b^{-1}a^{1/2} \leq I_{\mathcal{A}}.$$

Applying again Proposition 11.3.5, $b^{-1} \leq a^{-1}$. $\square$

**11.3.8. Corollary.**

*Let $\mathcal{A}$ be a unital $C^*$-algebra and $a \in \mathcal{A}$.*

*(i) If $a \in \mathcal{A}^+$, then $0 \leq a \leq \|a\| I_{\mathcal{A}}$, and $\sigma(a) \subset [0, \infty)$.*

*(ii) If $0 \leq a \leq b$, then $\|a\| \leq \|b\|$.*

*(iii) If $a \in \mathcal{A}^{sa}$, then $-\|a\| I_{\mathcal{A}} \leq a \leq \|a\| I_{\mathcal{A}}$, and $\sigma(a) \subset \mathbb{R}$.*

*Proof.* If $a \geq 0$, then $\sigma(a) \subset [0, \|a\|]$ by Lemma 11.3.2 and Theorem 9.2.12. Therefore $0 \leq a \leq \|a\| I_{\mathcal{A}}$ by Corollary 11.3.6. If now $0 \leq a \leq b$, from $b \leq \|b\| I_{\mathcal{A}}$ we get $a \leq \|b\| I_{\mathcal{A}}$. By Corollary 11.3.6 we then have $\|a\| = \operatorname{spr}(a) \leq \|b\|$.

When $a$ is selfadjoint, again by Proposition 9.2.24 and Proposition 11.2.1 we get that $\sigma(a) \subset \mathbb{R}$. We have $\sigma(a) \subset [-\|a\|, \|a\|]$ again by Theorem 9.2.12, so $-\|a\| I_{\mathcal{A}} \leq a \leq \|a\| I_{\mathcal{A}}$ by Corollary 11.3.6. $\square$

The set $\mathcal{A}^+ \subset \mathcal{A}$ is a positive cone, as we will see next. The “positive” part is trivial, since $a \geq 0$ and $a \leq 0$ imply that $\sigma(a) = \{0\}$ and then, as $a$ is normal, $\|a\| = \operatorname{spr}(a) = 0$.

***11.3.9. Proposition.***

*Let $a, b \in \mathcal{A}^+$, $t \geq 0$. Then $a + tb \in \mathcal{A}^+$.*

*Proof.* We have that $a + tb = (a + tb)^*$, so by Lemma 11.3.2 all we need to show is that $\sigma(a + tb) \subset [0, \infty)$. Since $a + tb$ is selfadjoint we have that $C^*(a + tb)$ is abelian. By Proposition 9.2.24,

$$\sigma(a + tb) = \{\tau(a + tb) : \ \tau \in \Sigma(C^*(a + b))\}.$$

As $\tau|_{C^*(a)} \in \Sigma(C^*(a))$ and $\tau|_{C^*(b)} \in \Sigma(C^*(b))$, we get from Proposition 9.2.24 that $\tau(a) \in \sigma(a)$ and $\tau(b) \in \sigma(b)$. So $\tau(a + tb) = \tau(a) + t\tau(b) \geq 0$ and then $\sigma(a + tb) \subset [0, \infty)$. □

We remark that if we just had the definition of positive as $a = c^*c$, $b = d^*d$ for some $c, d$, without the characterization of positive as "selfadjoint with non-negative spectrum" it seems like it would be really hard to show that there exists $z \in \mathcal{A}$ with $a + b = z^*z$.

The existence of the square root allows us to define the **absolute value** of $a$, namely $|a| = (a^*a)^{1/2} \in \mathcal{A}^+$.

***11.3.10. Proposition.***

*Let $a \in \mathcal{A}^{\mathrm{sa}}$. Then there exist unique $a_+, a_- \in \mathcal{A}^+$ such that $a = a_+ - a_-$ and $a_+a_- = 0$.*

*Proof.* Let

$$a_+ = \frac{|a| + a}{2}, \quad a_- = \frac{|a| - a}{2}.$$

Note that both $a_+, a_- \in C^*(a)$; in particular, they both commute with $a$. As $|\lambda| + \lambda \geq 0$ for all $\lambda \in \mathbb{R}$, Corollary 11.2.8 and Lemma 11.3.2 imply that $a_+, a_- \in \mathcal{A}^+$. Finally, since $|a|$ and $a$ commute (by continuous functional calculus, since $t|t| = |t|t$),

$$4a_+a_- = (|a| + a)(|a| - a) = |a|^2 - a^2 = a^2 - a^2 = 0.$$

It remains to address the uniqueness. Suppose that $a = a^*$ and $a = u - v$ with $u, v \geq 0$ and $uv = 0$. We have $a^2 = (u - v)^2 = u^2 + (-v)^2$, and more generally $a^n = u^n + (-v)^n$. So $p(a) = p(u) + p(-v)$ for all $p \in \mathbb{C}[x]$, and passing to the limit $f(a) = f(u) + f(-v)$ for all $f \in C_{\mathrm{zero}}(\sigma(a))$. If in particular $f(t) = \max\{t, 0\}$, then $f(u) = u$ and $f(-v) = 0$, since $u, v \geq 0$, and so

$$a_+ = f(a) = f(u) + f(-v) = f(u) = u,$$

and then $v = a_-$. □

We continue with a few examples of how the continuous functional calculus is used. The polar decomposition as in Proposition 10.4.11 makes no sense in an arbitrary C*-algebra, for $|a| \in \mathcal{A}$ for all $a$, but there need not be partial isometries in $\mathcal{A}$. For instance, in $C[0,1]$ the only partial isometries are 0 and 1, and in $C_0(\mathbb{R})$ the only partial isometry is 0. But we can still do a poor's man version of the polar decomposition by weakening the exponent:

***11.3.11. Proposition.***

*Let $\mathcal{A}$ be a C*-algebra, $a \in \mathcal{A}$, and $\alpha \in \left(0, \frac{1}{2}\right)$. Then there exists $u \in \mathcal{A}$ with $a = u(a^*a)^\alpha$, and $u^*u \leq (a^*a)^{1-2\alpha}$.*

*Proof.* The computations below require us to to work in the unitization $\tilde{\mathcal{A}}$, but the elements we obtain are in $\mathcal{A}$. Let $u_n = a\left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + a^*a\right]^{-1/2}(a^*a)^{1/2-\alpha} \in \mathcal{A}$. Let

$$x_{m,n} = \left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + a^*a\right]^{-1/2} - \left[\frac{1}{m}\, I_{\tilde{\mathcal{A}}} + a^*a\right]^{-1/2}.$$

Then, since $x_{m,n}$ commutes with $a^*a$ and its functional calculus

$$\begin{aligned}\|u_n - u_m\|^2 &= \|a x_{m,n}(a^*a)^{1/2-\alpha}\|^2 = \|(a^*a)^{1/2-\alpha} x_{m,n} a^*a x_{m,n}(a^*a)^{1/2-\alpha}\| \\ &= \|(a^*a)^{1/2} x_{m,n}(a^*a)^{1/2-\alpha}\|^2 = \|x_{m,n}\,(a^*a)^{1-\alpha}\|^2\end{aligned}$$

Consider the functions $f_n(t) = \left(\frac{1}{n} + t\right)^{-1/2} t^{1-\alpha}$. If $m > n$, then

$$f_m(t) - f_n(t) = t^{1-\alpha}\,\frac{\left(\frac{1}{n}+t\right)^{1/2} - \left(\frac{1}{m}+t\right)^{1/2}}{\left(\frac{1}{n}+t\right)^{1/2}\left(\frac{1}{m}+t\right)^{1/2}} \geq 0,$$

so the sequence $\{f_n\}$ is increasing on $[0,1]$, and it converges pointwise to $f(t) = t^{\frac{1}{2}-\alpha}$. By Dini's Theorem (2.6.15), the convergence is uniform. Hence

$$x_{m,n}(a^*a)^{1-\alpha} = f_m(a^*a) - f_n(a^*a) \to f(a^*a) - f(a^*a) = 0.$$

This proves that $\{u_n\}$ is Cauchy in $\mathcal{A}$, and so $u = \lim u_n \in \mathcal{A}$ exists. Now

$$u(a^*a)^\alpha = \lim_n u_n(a^*a)^\alpha = \lim_n a\left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + a^*a\right]^{-1/2}(a^*a)^{1/2} = a.$$

The justification of the last equality is as follows: if

$$\alpha = \|a - a\left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2}|a|\,\|^2,$$

$$\begin{aligned}\alpha &= \|a\Big(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2}|a|\Big)\|^2 \\ &= \Big\|\Big(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2}|a|\Big)a^*a\Big(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2}|a|\Big)\Big\|\end{aligned}$$

$$= \left\| \left(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2} |a|\right) |a|^2 \left(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2} |a|\right) \right\|$$
$$= \| |a| \left(I_{\tilde{\mathcal{A}}} - \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2} |a|\right) \|^2 = \|g_n(|a|)\|^2,$$

where
$$g_n(t) = |t| \left(1 - \left[\frac{1}{n} + |t|^2\right]^{-1/2} |t|\right) = |t| - \frac{|t|^2}{\left(\frac{1}{n} + |t|^2\right)^{1/2}}.$$

Another application of Dini's Theorem shows that $g_n \to 0$ uniformly on $[0, 1]$, and so $g_n(|a|) \to 0$.

Finally we have, using that everything commutes being functional calculus of $|a|$,

$$\begin{aligned} u_n^* u_n &= |a|^{1-2\alpha} \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2} a^* a \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1/2} |a|^{1-2\alpha} \\ &= |a|^{2-4\alpha} \, |a|^2 \left[\frac{1}{n} I_{\tilde{\mathcal{A}}} + |a|^2\right]^{-1} \\ &\le |a|^{2-4\alpha} = (a^* a)^{1-2\alpha}. \end{aligned}$$

The inequality will hold in the limit, so $u^* u \le (a^* a)^{1-2\alpha}$. □

***11.3.12. Proposition.***

*Let $\mathcal{A}$ be a non-unital $C^*$-algebra. Then the closed unit ball of $\mathcal{A}$ has no extreme points.*

*Proof.* Let $a \in \overline{B_1^{\mathcal{A}}(0)}$ be extreme. Suppose that $a^* a$ is not a projection (that is, $(a^* a)^2 \ne a^* a$). This implies that $|a|^{1/2} - |a| \ge 0$ is nonzero. By Proposition 11.3.11 we can write $a = u|a|^{1/2}$ for some $u \in \mathcal{A}$ with $\|u\| \le \|a\|^{1/4} \le 1$.

The condition $|a|^{1/2} \ne |a|$ forces $\sigma(|a|^{1/2}) \cap (0, 1) \ne \varnothing$. So there exists $t_0 \in \sigma(|a|^{1/2})$ with $0 < t_0 < 1$. Choose $\delta > 0$ so that $t_0 + \delta < 1$ and $t_0 - \delta > 0$. Let $c = \min\{t_0 - \delta, 1 - t_0 - \delta\}$ and choose $f \in C(\sigma(|a|^{1/2}))$ with $f(t_0) = c$, $f(t) \le f(t_0)$ for all $t$, and $f = 0$ outside of $(t_0 - \delta, t_0 + \delta)$. Then $0 \le t + f(t) \le 1$ for all $t \in [0, 1]$. Indeed, outside of $(t_0 - \delta, t_0 + \delta)$ we have $t + f(t) = t$, and when $t_0 - \delta \le t \le t_0 + \delta$ we have $t + f(t) \le t_0 + \delta + (1 - t_0 - \delta) = 1$. Let $b = f(|a|^{1/2})$, so that $|a|^{1/2} + b \le I_{\tilde{\mathcal{A}}}$; this guarantees that $|a|^{1/2} + b$ is in the closed unit ball of $\mathcal{A}$. We also have $t - f(t) \ge 0$ for $t - f(t) = t$ outside of $(t_0 - \delta, t_0 + \delta)$ and $t - f(t) \ge t_0 - \delta - c \ge 0$ when $t \in (t_0 - \delta, t_0 + \delta)$. Thus $0 \le |a|^{1/2} - b \le |a|^{1/2}$, showing that $|a|^{1/2} - b$ is also in the closed unit ball of $\mathcal{A}$.

Now we can write

$$a = u|a|^{1/2} = \frac{1}{2}\big(u(|a|^{1/2} + b\big) + \frac{1}{2}\big(u(|a|^{1/2} - b\big)$$

and $a$ would not be extreme. Thus $a^*a$ and similarly $aa^*$ are projections. Consider the projections $I_{\tilde{\mathcal{A}}} - a^*a$ and $I_{\tilde{\mathcal{A}}} - aa^*$ in $\tilde{\mathcal{A}}$. We claim that $(I_{\tilde{\mathcal{A}}} - aa^*)\mathcal{A}(I_{\tilde{\mathcal{A}}} - a^*a) = \{0\}$. For if that is not the case, there exists $b \in \mathcal{A}^+$ with $\|b\| = 1$ and $b = (I_{\tilde{\mathcal{A}}} - aa^*)c(I_{\tilde{\mathcal{A}}} - a^*a) = 0$. We have $ba^*a = 0$, so $ba^*ab^* = 0$, which gives $ba^* = 0$. So $b^*ba^*a = 0$ and similarly $aa^*bb^* = 0$. Hence

$$\|a + b\|^2 = \|(a^* + b^*)(a + b)\| = \|a^*a + b^*b\| = \max\{\|a^*a\|, \|b^*b\|\} = 1,$$

and similarly $\|a - b\| = 1$. This allows us to write $a = \frac{1}{2}(a + b) + \frac{1}{2}(a - b)$, contradicting that $a$ is extreme. This proves that $(I_{\tilde{\mathcal{A}}} - aa^*)\mathcal{A}(I_{\tilde{\mathcal{A}}} - a^*a) = \{0\}$. Given $b \in \mathcal{A}$, we have

$$(I_{\tilde{\mathcal{A}}} - aa^*)(I_{\tilde{\mathcal{A}}} - a^*a)b^*b(I_{\tilde{\mathcal{A}}} - a^*a)(I_{\tilde{\mathcal{A}}} - aa^*) = 0$$

(looking at the first four factors) and so $b(I_{\tilde{\mathcal{A}}} - a^*a)(I_{\tilde{\mathcal{A}}} - aa^*) = 0$ for all $b \in \mathcal{A}$. This implies that $(I_{\tilde{\mathcal{A}}} - a^*a)(I_{\tilde{\mathcal{A}}} - aa^*) = 0$ (see Exercise 11.1.3). This equality we can write as

$$I_{\tilde{\mathcal{A}}} = aa^* + a^*a - aa^*a^*a \in \mathcal{A},$$

showing that $\mathcal{A}$ is unital. Therefore, a non-unital C$^*$-algebra cannot have extreme points in the closed unit ball. □

***11.3.13. Proposition.***

*Let $\mathcal{A}$ be a unital C$^*$-algebra. Then the extreme points of the closed unit ball of $\mathcal{A}$ are precisely those $a \in \mathcal{A}$ such that $(I_{\tilde{\mathcal{A}}} - aa^*)\mathcal{A}(I_{\tilde{\mathcal{A}}} - a^*a) = \{0\}$. Such elements are necessarily partial isometries (that is, $a^*a$ and $aa^*$ are projections).*

*Proof.* Let $a \in \mathcal{A}$ be an extreme point of the closed unit ball. From the proof of Proposition 11.3.12 we know that $(I_{\mathcal{A}} - aa^*)\mathcal{A}(I_{\mathcal{A}} - a^*a) = \{0\}$. $(I_{\mathcal{A}} - aa^*)\mathcal{A}(I_{\mathcal{A}} - a^*a) = \{0\}$.

Conversely, if $(I_{\mathcal{A}} - aa^*)\mathcal{A}(I_{\mathcal{A}} - a^*a) = \{0\}$, we have in particular

$$0 = a^*(I_{\mathcal{A}} - aa^*)a(I_{\mathcal{A}} - a^*a) = a^*a(I_{\mathcal{A}} - a^*a)^2$$

Therefore $a^*a \geq 0$ and $\lambda \in \sigma(a^*a)$ satisfies $\lambda(1 - \lambda)^2 = 0$, so either $\lambda = 0$ or $\lambda = 1$. Then $a^*a - (a^*a)^2$ is normal with spectrum $\{0\}$ (Proposition 9.2.9), and so $(a^*a)^2 = a^*a$ by Corollary 11.3.8. So $a^*a$ is a projection; and as $\sigma(aa^*) \setminus \{0\} = \sigma(a^*a) \setminus \{0\}$ (Proposition 9.2.15) the same argument applies to show that $aa^*$ is also a projection. We have $aa^*a = a$, for $(aa^*a-a)^*(aa^*a-a) = 0$.

Now we write $p = a^*a$, $q = aa^*$ and suppose that $a = \frac{1}{2}(b+c)$ with $b, c$ in the closed unit ball. As $ap = a$, we also have $a = \frac{1}{2}(bp + cp)$, and

$$\begin{aligned} p = a^*a &= \frac{1}{4}\,(pb^* + pc^*)(bp + cp) = \frac{1}{4}\,(pb^*bp + pc^*cp + 2\mathrm{Re}\,pb^*cp) \\ &= \frac{1}{4}\,(pb^*bp + pc^*cp + p(2\mathrm{Re}\,b^*c)p) \leq \frac{1}{4}\,(pb^*bp + pc^*cp + 2p) \leq p. \end{aligned} \tag{11.7}$$

It follows that $2p \leq pb^*bp + pc^*cp \leq 2p$. Thus $0 = (p - pb^*bp) + (p - pc^*cp)$ is a sum of two positive operators, and so $p = pb^*bp = pc^*cp$. But if now we go to the equalities (11.7) we get $p = \frac{1}{2}\,(pb^*cp + pc^*bp)$. As $p$ is extreme in $p\mathcal{A}p$ by Exercise 11.3.2, we conclude that $pb^*cp = pc^*bp = p$. Now

$$p(b-c)^*(b-c)p = pb^*bp + pc^*cp - pc^*bp - pb^*cp = 2p - 2p = 0.$$

So $bp = cp$. Now we can use that $qa = q$ and develop an analog argument as before but with $q$ on the left instead of $p$ on the right. Thus $qb = qc$. From this we have

$$b - c = (I_{\mathcal{A}} - q)(b-c)(I_{\mathcal{A}} - p) = 0$$

and therefore $a$ is an extreme point. $\square$

**11.3.1. Exercises.**

(11.3.1) Let $a \in \mathcal{A}$. Show that $\|\operatorname{Re} a\| \leq \|a\|$ and $\|\operatorname{Im} a\| \leq \|a\|$.

(11.3.2) Let $\mathcal{A}$ be a C$^*$-algebra. Show that $I_{\mathcal{A}}$ is an extreme point in the closed unit ball of $\mathcal{A}$.

(11.3.3) Let $a \in \mathcal{A}^+$. Show that its positive square root is unique.

(11.3.4) Let $\mathcal{A}$ be a C$^*$-algebra and $a, b \in \mathcal{A}^+$. Show that if $b \leq a$, then $b^{1/2} \leq a^{1/2}$.

(11.3.5) Show by example that it is not true in general that if $0 \leq b \leq a$, then $b^2 \leq a^2$.

(11.3.6) Let $a \in \mathcal{A}$ be selfadjoint. Show that the following statements are equivalent:

(i) $a \geq 0$;

(ii) $\|\gamma\, I_{\mathcal{A}} - a\| \leq \gamma$ for all $\gamma \geq \|a\|$;

(iii) there exists $\gamma \in \mathbb{R}$ with $\gamma \geq \|a\|$ and $\|\gamma\, I_{\mathcal{A}} - a\| \leq \gamma$.

(11.3.7) Let $\mathcal{A}$ be a C$^*$-algebra. A function $f : [0, \infty) \to [0, \infty)$ is **operator monotone** if $f(a) \leq f(b)$ whenever $0 \leq a \leq b$.

(i) Use Corollary 11.3.7 to prove that $f_s(t) = st(1 + st)^{-1}$ is operator monotone for each $s > 0$.

(ii) Prove that $g(t) = t^\beta$ is operator monotone for all $\beta \in (0, 1)$ by considering $g(t) = \int_0^\infty f_s(t) s^{-\beta-1}\, ds$.

(11.3.8) Let $\mathcal{A}$ be a C$^*$-algebra and $a, p \in \mathcal{A}$ with $a \geq 0$ and $p$ a projection ($p = p^2 = p^*$). Show that $a \leq p$ if and only if $\|a\| \leq 1$ and $a = pa = ap$. *(When $a$ and $p$ are projections in $\mathcal{B}(\mathcal{H})$ this was done in Proposition 10.5.3; now we need rather similar arguments, but they have to be entirely algebraic)*

(11.3.9) Let $\mathcal{A}$ be a C$^*$-algebra and $a \in \mathcal{A}$ nonzero. Show that there exist $r \in \{1, -1\}$ and $s \in \{1, i\}$ such that $(\operatorname{Re} rsa)^+ \neq 0$.

(11.3.10) Let $\pi : \mathcal{A} \to \mathcal{B}$ be a $*$-homomorphism. This is an alternative way to prove that an injective $*$-homomorphism is isometric, that was proven in Proposition 11.2.10.

(i) Show that $\pi(a) \geq 0$ if and only if $a \geq 0$.

(ii) Conclude that if $\pi$ is injective, then it is isometric.

## 11.4. Ideals and $*$-Homomorphisms

By an **ideal** $\mathcal{J}$ of $\mathcal{A}$, we mean a bilateral and closed ideal. When $\mathcal{A} = C(X)$ for a compact Hausdorff space $X$, the ideals of $\mathcal{A}$ were characterized in Proposition 7.4.13. Given an ideal $\mathcal{J} \subset \mathcal{A}$, the quotient $\mathcal{A}/\mathcal{J}$ is a Banach space as in Section 5.3. The fact that $\mathcal{J}$ is an ideal allows us to immediately give $\mathcal{A}/\mathcal{J}$ a multiplicative structure, by defining $(a+\mathcal{J})(b+\mathcal{J}) = ab+\mathcal{J}$. It is less obvious that $\mathcal{J}$ contains adjoints, and that the norm is a Banach algebra norm, and a C$^*$-norm; that will require approximate units.

**11.4.1. Definition.**

*Let $\mathcal{A}$ be a C$^*$-algebra. An **approximate unit** is an increasing net $\{e_j\} \subset \mathcal{A}^+$ such that $0 \leq e_j \leq I_{\mathcal{A}}$ and $\lim_j ae_j = a$ for all $a \in \mathcal{A}$.*

Because $\mathcal{A}$ contains adjoints and the adjoint operator is continuous, we get that

$$\forall a \in \mathcal{A},\ \lim_j ae_j = a \iff \forall a \in \mathcal{A},\ \lim_j e_j a = a. \tag{11.8}$$

*11.4.2. Examples.* Here are some canonical examples of approximate units in non-unital C$^*$-algebras. In all cases they are countable, although that will not be the case in general. We note that in these examples we are **not** using $e_n$ to denote the canonical basis.

(i) $\mathcal{A} = c_0$. Take $e_n = \sum_{j=1}^n \delta_j$, i.e. $e_n = (\overbrace{1,\ldots,1}^{n},0,\ldots)$. Then $0 \leq e_n \leq e_{n+1} \leq 1$ for all $n$, and $\lim_n e_n a = a$ for all $a \in c_0$.

(ii) $\mathcal{A} = C_0(\mathbb{R})$. Take $e_n$ to be a continuous function that is 1 on $[-n,n]$ and then decreases to zero for $|x| \geq n+1$. Then $0 \leq e_n \leq e_{n+1} \leq 1$ for all $n$, and $\lim_n e_n f = f$ for all $f \in C_0(\mathbb{R})$.

(iii) $\mathcal{A} = \mathcal{K}(\mathcal{H})$. Fix an orthonormal basis $\{\xi_n\}$ of $\mathcal{H}$ and let $E_n = \sum_{k=1}^n P_k$, where $P_k = \langle \cdot, \xi_k \rangle \xi_k$ is the orthogonal projection onto $\mathbb{C}\xi_k$. Then we have $\lim_n E_n T = T$ for all $T \in \mathcal{K}(\mathcal{H})$ (see the proof of Proposition 10.6.2). Note that this is *not* an approximate unit for $\mathcal{B}(\mathcal{H})$.

***11.4.3. Lemma.***

*Let $a, b \in \mathcal{A}^+$ with $a \leq b$. Then $0 \leq a(I_{\mathcal{A}}+a)^{-1} \leq b(I_{\mathcal{A}}+b)^{-1}$.*

*Proof.* The inequality $a(I_{\mathcal{A}}+a)^{-1} \leq b(I_{\mathcal{A}}+b)^{-1}$ was shown in Exercise 11.3.7. We also have, by Corollary 11.2.8, that $a(I_{\mathcal{A}}+a)^{-1} \geq 0$. ☐

***11.4.4. Theorem.***

*Let $\mathcal{A}$ be a $C^*$-algebra. Then $\mathcal{A}$ has an approximate unit.*

*Proof.* Let $\mathcal{Z} = \{a \in \mathcal{A}^+ : \|a\| < 1\}$; make it a net by using the partial order $\leq$. This requires us to show that if $a, b \in \mathcal{Z}$, there exists $c \in \mathcal{Z}$ with $a \leq c$ and $b \leq c$. Let $a' = a(I_{\mathcal{A}}-a)^{-1}$, $b' = b(I_{\mathcal{A}}-b)^{-1}$, and $c = (a'+b')(I_{\mathcal{A}}+a'+b')^{-1}$. From $\|a\| < 1$ and Corollary 11.2.8 and Lemma 11.3.2 we get that $a', b' \in \mathcal{A}^+$. Let $\beta = \|a'+b'\|$. Since $a'+b' \leq \beta\, I_{\mathcal{A}}$ (Corollary 11.3.8), from Lemma 11.4.3 we have $c = (a'+b')(I_{\mathcal{A}}+a'+b')^{-1} \leq \frac{\beta}{1+\beta} I_{\mathcal{A}}$, and then by Corollary 11.3.6 $\|c\| = \operatorname{spr}(c) \leq \frac{\beta}{1+\beta} < 1$. Thus $c \in \mathcal{Z}$. From $a' \leq a'+c'$ and Lemma 11.4.3,

$$a = a'(I_{\mathcal{A}}+a')^{-1} \leq (a'+b')(I_{\mathcal{A}}+a'+b')^{-1} = c.$$

Similarly, $b \leq c$. So $\mathcal{Z}$ is directed.

By Proposition 11.3.10 we have that $\mathcal{A} = \operatorname{span} \mathcal{Z}$. So it is enough to show that $\lim_z za = a$ for $a \in \mathcal{Z}$. Fix $a \in \mathcal{Z}$ and $\varepsilon > 0$. By Theorem 11.2.3, we have a $*$-isomorphism $\Gamma$ between $C^*(a)$ and $C_0(T)$, where $T = \Sigma(C^*(a))$. Let $f = \Gamma(a)$ be the image of $a$ under the isomorphism, and let $K = \{w \in T : |f(w)| \geq \varepsilon\}$. Since $f$ is continuous and bounded, $K$ is compact. By Urysohn's Lemma (Theorem 2.6.5) there exists $g : T \to [0,1]$ continuous, with compact support, and such that $g|_K = 1$, $0 \leq g \leq 1$. Fix $\delta > 0$ with $\delta < 1$ and $0 < 1-\delta < \varepsilon$. Then $\delta g \in \mathcal{Z}$ and $\|f - \delta g f\| \leq \varepsilon$; indeed, if $w \in K$ then $|f(w) - \delta g(w) f(w)| = |f(w)|\,(1-\delta) < \varepsilon$ (note that $\|f\| = \|a\| < 1$), and if $w \notin K$ then $|f(w) - \delta g(w) f(w)| \leq |f(w)| < \varepsilon$. So if $z = \Gamma^{-1}(\delta g)$, we have $\|z\| = \|\delta g\| \leq \delta < 1$, so $z \in \mathcal{Z}$, and $\|a - za\| \leq \varepsilon$.

If now $z' \geq z$ in $\mathcal{Z}$, on the unitization $\tilde{\mathcal{A}}$ we have $I_{\mathcal{A}} - z' \leq I_{\mathcal{A}} - z$, and so $a(I_{\mathcal{A}}-z')a \leq a(I_{\mathcal{A}}-z)a$. Then, using Corollary 11.3.8 for the second to last inequality,

$$\begin{aligned}
\|a - z'a\|^2 &= \|(I_{\mathcal{A}}-z')^{1/2}(I_{\mathcal{A}}-z')^{1/2}a\|^2 \\
&\leq \|(I_{\mathcal{A}}-z')^{1/2}a\|^2 = \|a(I_{\mathcal{A}}-z')a\| \\
&\leq \|a(I_{\mathcal{A}}-z)a\| \leq \|(I_{\mathcal{A}}-z)a\| = \|a - za\| \leq \varepsilon.
\end{aligned}$$

Thus $\lim_z \|a - za\| = 0$. Though the limit was obtained in the unitization of $\mathcal{A}$, it is about elements of $\mathcal{A}$, and the norm of $\mathcal{A}$ agrees with its norm as a subalgebra of the unitization. □

When $\mathcal{A}$ is separable, one can obtain a countable approximate unit (Exercise 11.4.3). For now our main use of approximate units will be when dealing with ideals. We start with a technical result. As of now we don't know that ideals are C*-subalgebras (because in the definition we don't require them to contain adjoints), so Proposition 11.4.5 below does not follow directly from Theorem 11.4.4.

We will again use the identity in the proof below. This does not require for $\mathcal{A}$ to be unital, but rather we can work in the unitization, and elements of the form $a(I_{\mathcal{A}} - e_j)$, though expressed in $\tilde{\mathcal{A}}$, are in $\mathcal{A}$.

**11.4.5. Proposition.**

*Let $\mathcal{J} \subset \mathcal{A}$ be an ideal. Then there exists an increasing net $\{e_j\} \subset \mathcal{J} \cap \mathcal{A}^+$, such that $\lim_j \|a - ae_j\| = 0$ for all $a \in \mathcal{J}$. As a consequence, $\mathcal{J} = \mathcal{J}^*$. In particular, ideals are C*-subalgebras.*

*Proof.* Consider $\mathcal{B} = \mathcal{J} \cap \mathcal{J}^*$. Then $\mathcal{B} \subset \mathcal{A}$ and it is a C*-algebra. By Theorem 11.4.4, there exists an approximate unit $\{e_j\} \subset \mathcal{B}$. Now, for $a \in \mathcal{J}$, working in $\tilde{\mathcal{A}}$ but taking into account that all the products are in $\mathcal{A}$ anyway,

$$\begin{aligned}\|a - ae_j\|^2 &= \|(a(I_{\mathcal{A}} - e_j)\|^2 = \|(I_{\mathcal{A}} - e_j)a^*a(I_{\mathcal{A}} - e_j)\| \\ &\leq \|a^*a(I_{\mathcal{A}} - e_j)\| = \|a^*a - a^*ae_j\| \to 0,\end{aligned}$$

as $a^*a \in \mathcal{B}$. Now take any $a \in \mathcal{J}$. Since $a = \lim_j ae_j$, we get $a^* = \lim e_j a^* \in \mathcal{J}$. □

The C*-algebra $\mathcal{B} = \mathcal{J} \cap \mathcal{J}^*$ is never trivial if $\mathcal{J} \neq \{0\}$. Indeed, if $a \in \mathcal{J}$ with $a \neq 0$ then $a^*a \in \mathcal{B}$ is nonzero.

**11.4.6. Definition.**

*The C*-algebra $\mathcal{A}$ is **simple** if its only ideals are $\{0\}$ and $\{\mathcal{A}\}$ (note that we are always talking about closed bilateral ideals here).*

*The **centre** of $\mathcal{A}$ is the set*

$$\mathcal{Z}(\mathcal{A}) = \{a \in \mathcal{A} : \ ab = ba, \ b \in \mathcal{A}\}.$$

It is not hard to check that $\mathcal{Z}(\mathcal{A})$ is an abelian C*-subalgebra of $\mathcal{A}$ (Exercise 11.4.1).

***11.4.7. Proposition.***

*Let $\mathcal{A}$ be a simple C*-algebra. Then $\mathcal{Z}(\mathcal{A}) = \mathbb{C}\, I_{\mathcal{A}}$ if $\mathcal{A}$ is unital, and $\mathcal{Z}(\mathcal{A}) = \{0\}$ if $\mathcal{A}$ is not unital.*

*Proof.* We consider separately the unital and non-unital cases. Assume first that $\mathcal{A}$ is unital. Choose $a \in \mathcal{Z}(\mathcal{A})$ nonzero and pick $\lambda \in \sigma(a)$. Consider $\mathcal{J} = \overline{(a - \lambda I_{\mathcal{A}})\mathcal{A}}$. Because $a$ is central, $\mathcal{J}$ is an ideal. Because $a - \lambda I_{\mathcal{A}}$ is not invertible, for any $b \in \mathcal{A}$ we have that $(a - \lambda I_{\mathcal{A}})b$ is not invertible (using again that $a$ is central). By Proposition 6.2.3,

$$\|(a - \lambda I_{\mathcal{A}})b - I_{\mathcal{A}}\| \geq 1, \qquad b \in \mathcal{A}.$$

This shows that $I_{\mathcal{A}}$ is at distance 1 from $\mathcal{J}$. As $\mathcal{A}$ is simple, the only possibility is that $\mathcal{J} = 0$, which then implies that $a = \lambda\, I_{\mathcal{A}}$.

If we assume that $\mathcal{A}$ is non-unital, again choose $a \in \mathcal{Z}(\mathcal{A})$ nonzero. Since $a^*a \in \mathcal{Z}(\mathcal{A})$, we may assume that $a$ is positive; because $a$ is selfadjoint and nonzero, its spectrum cannot be $\{0\}$. We want to do some functional calculus on $\mathcal{Z}(\mathcal{A})$; since it is non-unital, $0 \in \sigma(a)$ and only continuous functions $f$ with $f(0) = 0$ are allowed. If distinct nonzero $\lambda_1, \lambda_2 \in \sigma(a)$ exist, consider $g_1, g_2 \in C(\sigma(a))$ with $g_1(\lambda_1) = g_2(\lambda_2) = 1$ and $g_1 g_2 = 0$. Then the elements $b_1 = g_1(a)$ and $b_2 = g_2(a)$ are central, nonzero, and $b_1 b_2 = 0$. Because each is central, $b_1\mathcal{A}$ and $b_2\mathcal{A}$ are distinct nonzero ideals (note that their product is zero), so at least one of them is not $\mathcal{A}$ and this contradicts the simplicity of $\mathcal{A}$. This shows that $\sigma(a)$ can have at most a single nonzero element. So $\sigma(a) = \{0, \lambda\}$. Let $p = \lambda^{-1}a$. This element is positive with $\sigma(p) = \{0, 1\}$; hence $p^2 = p$ by Corollary 11.3.8, since $\sigma(p^2 - p) = \{0\}$. As $p \in \mathcal{Z}(\mathcal{A})$, we can consider the ideal $p\mathcal{A}$. It is nonzero, and since $\mathcal{A}$ is simple then $p\mathcal{A} = \mathcal{A}$. Now any $b \in \mathcal{A}$ is of the form $b = pc$ for some $c$. Then $pb = p^2c = pc = b$; as $p$ is central, this implies that $p$ is a unit for $\mathcal{A}$, a contradiction. □

The converse to Proposition 11.4.7 is false, with the easiest example being $\mathcal{B}(\mathcal{H})$ for an infinite-dimensional Hilbert space $\mathcal{H}$; the centre is trivial, but $\mathcal{K}(\mathcal{H})$ is a proper ideal.

We will now talk about quotients. We dealt with quotients of Banach spaces (which C*-algebras are) in Section 5.3.

***11.4.8. Theorem.***

*Let $\mathcal{J} \subset \mathcal{A}$ be an ideal. Then $\mathcal{A}/\mathcal{J}$ is a C*-algebra.*

*Proof.* Let $\{e_j\}$ be an approximate unit of $\mathcal{J}$ as in Proposition 11.4.5. We will show first that

$$\|a+\mathcal{J}\| = \lim_j \|a - e_j a\| = \lim_j \|a - a e_j\|. \tag{11.9}$$

Of course, it is enough to show one of the two equalities since the other follows from taking adjoints. We have $\|a - e_j a\| \geq \|a + \mathcal{J}\|$ by definition of the quotient norm. Now fix $\varepsilon > 0$, and choose $b \in \mathcal{J}$ such that $\|a - b\| < \|a + \mathcal{J}\| + \varepsilon/2$. As $b = \lim_j e_j b$, there exists $j_0$ such that $\|b - e_j b\| < \varepsilon/2$ for all $j \geq j_0$. Then, for all $j \geq j_0$, and working on $\tilde{\mathcal{A}}$,

$$\begin{aligned}\|a - e_j a\| &\leq \|a - b + e_j b - e_j a\| + \|b - e_j b\| \\ &= \|(I_{\mathcal{A}} - e_j)(a-b)\| + \|b - e_j b\| \leq \|a-b\| + \|b - e_j b\| \\ &< \|a + \mathcal{J}\| + \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \|a + \mathcal{J}\| + \varepsilon.\end{aligned}$$

Combining both estimates we get that $\|a+\mathcal{J}\| = \lim_j \|a - e_j a\|$, thus proving (11.9). In particular we get that

$$\|(a+\mathcal{J})^*\| = \|a^* + \mathcal{J}\| = \lim_j \|a^* - e_j a^*\| = \lim_j \|a - a e_j\| = \|a + \mathcal{J}\|.$$

We can use (11.9) to show that the quotient norm is a Banach algebra norm (that is, submultiplicative). Indeed, using that $\mathcal{J}$ is an ideal

$$\begin{aligned}\|ab + \mathcal{J}\| &\leq \|ab + a e_j b e_j - a e_j b - a b e_j\| = \|(a - a e_j)(b - b e_j)\| \\ &\leq \|a - a e_j\| \, \|b - b e_j\|;\end{aligned}$$

as we can do the above for all $j$, taking limit we get $\|ab+\mathcal{J}\| \leq \|a+\mathcal{J}\| \, \|b+\mathcal{J}\|$.

Now, using (11.9) twice and working again on the unitization,

$$\begin{aligned}\|a + \mathcal{J}\|^2 &= \lim_j \|a - a e_j\|^2 = \lim_j \|a(I_{\mathcal{A}} - e_j)\|^2 \\ &= \lim_j \|(I_{\mathcal{A}} - e_j) a^* a (I_{\mathcal{A}} - e_j\| \\ &\leq \lim_j \|a^* a (I_{\mathcal{A}} - e_j)\| = \lim_j \|a^* a - a^* a e_j\| = \|a^* a + \mathcal{J}\|.\end{aligned}$$

And $\|a^*a+\mathcal{J}\| \leq \|a^*+\mathcal{J}\| \, \|a+\mathcal{J}\| = \|a+\mathcal{J}\|^2$, establishing the equality. $\square$

It turns out that the infimum in the definition of the norm $\|a+\mathcal{J}\|$ is always a minimum in the case of C$^*$-algebras; that is, there always exists $b \in \mathcal{J}$ such that $\|a + \mathcal{J}\| = \|a - b\|$; this is a consequence of Proposition 11.7.6.

With quotients at our disposal, we can now complete our journey towards the basic properties of ∗-homomorphisms.

***11.4.9. Proposition.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then $\pi$ is bounded, $\|\pi\| \leq 1$, and $\pi(\mathcal{A})$ is closed. If $\pi$ is injective, then $\pi$ is isometric.*

*Proof.* In light of Propositions 11.2.10 and 11.1.11, it remains to show that $\pi(\mathcal{A})$ is closed.

Suppose first that $\pi$ is injective; then we know that it is isometric. If $\{\pi(a_j)\} \subset \pi(\mathcal{A})$ is Cauchy, then $\|a_j - a_k\| = \|\pi(a_j - a_k)\| = \|\pi(a_j) - \pi(a_k)\|$, so $\{a_j\}$ is Cauchy; then by the completeness of $\mathcal{A}$ there exists $a = \lim a_j \in \mathcal{A}$ and $\pi(a) = \lim \pi(a_j)$, showing that $\pi(\mathcal{A})$ is closed. In the case where $\pi$ is not injective, since $\pi$ is bounded $\ker \pi$ is a closed ideal. If we now define $\tilde{\pi} : \mathcal{A}/\ker \pi \to \overline{\pi(\mathcal{A})}$ by $\tilde{\pi}(a + \ker \pi) = \pi(a)$, we have that $\tilde{\pi}$ is an injective $*$-homomorphism. By the above reasoning, $\pi(\mathcal{A}) = \tilde{\pi}(\mathcal{A}/\ker \pi)$ is closed. ☐

The result in Proposition 11.4.9 does not hold if we weaken the assumption so that $\mathcal{A}$ is a dense $*$-subalgebra of a C*-algebra (Exercise 11.4.6). Nor does it work for arbitrary Banach algebras. Indeed, recall that $\ell^p(\mathbb{N})$ is a Banach algebra—with pointwise multiplication—for $1 \leq p \leq \infty$ (Examples 9.2.3). For $p < p' < \infty$ we may consider the identity map $\iota : \ell^p(\mathbb{N}) \to \ell^{p'}(\mathbb{N})$. The map $\iota$ is a Banach algebra monomorphism, but its range is not closed, and hence it cannot isometric (see Theorem 9.8.16).

We record here the fact, used in the proof of Proposition 11.4.9, that we have a First Isomorphism Theorem for $C^*$-algebras:

***11.4.10. Theorem.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then there is an isomorphism $\mathcal{A}/\ker \pi \simeq \pi(\mathcal{A})$. In particular, $\pi(\mathcal{A})$ is a $C^*$-subalgebra of $\mathcal{B}$.*

### 11.4.1. The Calkin Algebra.

**11.4.11. Definition.**

*The **Calkin Algebra** is the quotient algebra $\mathcal{C}(\mathcal{H}) = \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$.*

Very easy to define, but not so easy to understand. When we mod $\mathcal{B}(\mathcal{H})$ by the compacts, we are somehow "erasing" everything about $\mathcal{B}(\mathcal{H})$ that is finite-dimensional. Here is an example of what we mean: if we write $\pi : \mathcal{B}(\mathcal{H}) \to$

$\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ the quotient map, and we choose $\varphi \in \mathcal{C}(\mathcal{H})^*$, we can induce $\tilde{\varphi} \in \mathcal{B}(\mathcal{H})^*$ by $\tilde{\varphi} = \varphi \circ \pi$. This functional $\varphi$ on $\mathcal{B}(\mathcal{H})$, is zero on all compact operators; in particular $\tilde{\varphi}(E_{kj}) = 0$ for all $k, j$ and any set $\{E_{kj}\}$ of matrix units.

***11.4.12. Lemma.***

*There exists an uncountable family $\{A_t\}_{t\in(0,1)}$ of infinite subsets of $\mathbb{N}$ such that, for all $s, t \in (0,1)$, $A_s \cap A_t$ is finite if $s \neq t$.*

*Proof.* Given $t \in (0,1)$, write $t = \sum_{k=1}^{\infty} b_k\, 10^{-k}$ its decimal expansion. Let

$$A_t = \Big\{ \sum_{k=1}^{m} b_k\, 10^{m-k} \,:\, m \in \mathbb{N} \Big\}.$$

So for instance, if $t = 0.235$, then $A_t = \{2, 23, 235, 2350, 23500, \ldots\}$. If $s \neq t$, then their expansions differ at some point, giving us that $A_s \cap A_t$ is finite. □

The point of Lemma 11.4.12 is to be used in Proposition 11.4.13 below, which shows that $\mathcal{C}(\mathcal{H})$ is an example of a C$^*$-algebra that cannot act on a separable Hilbert space.

***11.4.13. Proposition.***

*Let $\mathcal{H}$ be an infinite-dimensional Hilbert space and $\pi : \mathcal{C}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be any faithful (that is, injective) representation. Then $\mathcal{K}$ is non-separable.*

*Proof.* We will prove that $\mathcal{C}(\mathcal{H})$ has an uncountable family of pairwise orthogonal projections; this establishes the non-separability of $\mathcal{K}$ as such projections form an uncountable set where any two elements are at fixed distance (or equivalently, it allows one to create an uncountable orthonormal basis).

Consider an uncountable family $\{A_t\}$ of subsets of $\mathbb{N}$ with finite intersections, as in Lemma 11.4.12. Let $\{E_{kj}\}$ be the usual matrix units in $\mathcal{B}(\mathcal{H})$; if $\mathcal{B}(\mathcal{H})$ is not separable, we only construct a countable set of matrix units (that is, instead of using a full orthonormal basis we use a countable orthonormal set). Define projections

$$P_t = \sum_{k\in A_t} E_{kk}.$$

So we get uncountable many projections, with

$$P_t P_s = \sum_{k\in A_t\cap A_s} E_{kk}$$

of finite-rank. Then $\pi(P_t)\pi(P_s) = \pi(P_tP_s) = 0$, and the family $\{\pi(P_t)\}_{t\in(0,1)}$ is an uncountable family of pairwise orthogonal projections in $\mathcal{C}(\mathcal{H})$. $\square$

**11.4.2. Exercises.**

(11.4.1) Let $\mathcal{A}$ be a C*-algebra. Show that $\mathcal{Z}(\mathcal{A})$ is a C*-subalgebra of $\mathcal{A}$.

(11.4.2) Show the equivalence (11.8).

(11.4.3) If $\mathcal{A}$ is separable, show that it admits a countable approximate unit *(Hint: consider an increasing sequence of finite sets with dense union)*

(11.4.4) Let $\mathcal{A}$ be a C*-algebra and $\{e_j\} \subset \mathcal{A}$ an approximate unit. Show that $\lim_j \|e_j\| = \lim_j \|e_j^2\| = 1$.

(11.4.5) Let $\mathcal{B} \subset \mathcal{A}$ be an inclusion of C*-algebras. Suppose that $\mathcal{B}$ has an approximate unit $\{e_j\}$ that is also an approximate unit for $\mathcal{A}$. Show that any other approximate unit of $\mathcal{B}$ is an approximate unit for $\mathcal{A}$.

(11.4.6) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{A}_0 \subset \mathcal{A}$ a dense $*$-subalgebra. Let $\mathcal{B}$ be a C*-algebra and $\pi : \mathcal{A}_0 \to \mathcal{B}$ an injective $*$-homomorphism. Show by example that $\pi$ is not necessarily bounded.

(11.4.7) Let $\mathcal{A}$ be approximately finite dimensional (AFD); that is, $\mathcal{A} = \overline{\bigcup_n \mathcal{A}_n}$, where $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ and $\dim \mathcal{A}_n < \infty$ for all $n$. Show that $\mathcal{A}$ has a countable approximate unit made out of projections. *(Hint: use that finite-dimensional C*-algebras are unital, Lemma 11.8.2)*

(11.4.8) Let $\mathcal{A}$ be a C*-algebra and $\{e_s\}$ an approximate unit for $\mathcal{A}$. Consider linearly independent elements $b_1, \ldots, b_m \in \mathcal{A}$. Show that there exists $s_0$ such that for all $s \geq s_0$, $e_sb_1, \ldots, e_sb_m$ are linearly independent.

## 11.5. States

We have already used duals extensively. When considering C*-algebras, the order on $\mathcal{A}$ induces an order on $\mathcal{A}^*$.

**11.5.1. Definition.**

*Let $\mathcal{A}$ be a C*-algebra. We say that $\varphi \in \mathcal{A}^*$ is* ***positive*** *if $\varphi(a) \geq 0$ whenever $a \geq 0$. The* ***state space*** *of $\mathcal{A}$ is the set*

$$S(\mathcal{A}) = \{\varphi \in \mathcal{A}^* : \ \|\varphi\| = 1 \ \text{ and } \varphi \geq 0\}.$$

*A linear functional $\phi \in \mathcal{A}^*$ is* ***Hermitian*** *if $\phi(a) \in \mathbb{R}$ for all selfadjoint $a \in \mathcal{A}$.*

A straightforward observation is that $S(\mathcal{A})$ is convex. It is also routine to check that Hermitian functionals are precisely those $\phi \in \mathcal{A}$ such that $\phi(a^*) = \overline{\phi(a)}$ for all $a \in \mathcal{A}$ (Exercise 11.5.1). Positive linear functionals are Hermitian (Exercise 11.5.3).

An everyday tool for dealing with positive linear functionals is the already familiar **Cauchy–Schwarz inequality**:

***11.5.2. Theorem. (Cauchy–Schwarz)***

*Let $\varphi : \mathcal{A} \to \mathbb{C}$ be linear and positive. Then*

$$|\varphi(b^*a)| \leq \varphi(a^*a)^{1/2}\varphi(b^*b)^{1/2} \qquad a, b \in \mathcal{A}. \tag{11.10}$$

*In particular we have the **Kadison–Schwarz inequality**: if $\varphi$ is bounded (which it always is, see Proposition 11.5.4),*

$$|\varphi(a)|^2 \leq \|\varphi\| \, \varphi(a^*a), \qquad a \in \mathcal{A}. \tag{11.11}$$

*Proof.* The proof of (11.10) is exactly that of Theorem 4.2.2, applied to the sesquilinear positive form $(a, b) \longmapsto \varphi(b^*a)$.

As for (11.11), if $\mathcal{A}$ is unital, from (11.10) we get

$$|\varphi(a)|^2 = |\varphi(I_{\mathcal{A}}a)|^2 \leq \varphi(I_{\mathcal{A}}^* I_{\mathcal{A}})\varphi(a^*a) = \varphi(I_{\mathcal{A}})\varphi(a^*a) \leq \|\varphi\| \, \varphi(a^*a).$$

In the non-unital case, from the Cauchy–Schwarz inequality we have for any $b \in \mathcal{A}$ with $\|b\| \leq 1$

$$|\varphi(ba)|^2 \leq \varphi(bb^*)\varphi(a^*a) \leq \|\varphi\| \, \varphi(a^*a).$$

By Theorem 11.4.4 there exists an approximate unit in $\mathcal{A}$, which implies that there exists a sequence $\{b_n\} \subset \mathcal{A}$ such that $\|b_n\| \leq 1$ and $b_n a \to a$, which finishes the proof via the continuity of $\varphi$. $\square$

*11.5.3. Remark.* In the unital case, a nice direct proof of (11.11) for states is as follows: if $\varphi$ is a state

$$0 \leq \varphi(|a - \varphi(a)\,1|^2) = \varphi(a^*a + |\varphi(a)|^2 - 2\mathrm{Re}\, a\,\overline{\varphi(a)}) = \varphi(a^*a) - |\varphi(a)|^2.$$

When $\varphi$ is a positive linear functional, the above argument can be applied to $\varphi/\|\varphi\|$ after we know—via Proposition 11.5.4—that when we normalize a positive linear functional it becomes unital. This idea also works in the non-unital case by using an approximate unit and the fact—to be proven in Proposition 11.5.4—that $\|\varphi\| = \lim_j \varphi(e_j)$.

Since we consider our linear functionals already in the dual of $\mathcal{A}$, they are assumed to be bounded. But it is interesting to note that when $\varphi \geq 0$, it is automatically bounded (in particular, boundedness of $\varphi$ was not a necessary requirement in Theorem 11.5.2); we include this fact in the following characterization of positivity.

In Proposition 11.5.4 below, we note that when $\mathcal{A}$ is unital $\{I_{\mathcal{A}}\}$ is an approximate identity for $\mathcal{A}$, so the condition (11.12) can be written as $\|\varphi\| = \varphi(I_{\mathcal{A}})$.

***11.5.4. Proposition.***

*Let $\varphi : \mathcal{A} \to \mathbb{C}$ be linear. The following statements are equivalent:*

*(i)* $\varphi \geq 0$;

*(ii)* $\varphi$ *is bounded and for any approximate unit* $\{e_j\}$ *in* $\mathcal{A}$ *we have*

$$\|\varphi\| = \lim_j \varphi(e_j). \tag{11.12}$$

*The equality* (11.12) *becomes* $\|\varphi\| = \varphi(I_{\mathcal{A}})$ *when* $\mathcal{A}$ *is unital.*

*Proof.* (i) $\Longrightarrow$ (ii): we first show that $\varphi$ is bounded. Suppose it is not. Because every element in $\mathcal{A}$ with norm at most 1 can be written as $a = a_1 - a_2 + i(a_3 - a_4)$, with $a_j \geq 0$ and $\|a_j\| \leq 1$, $j = 1, 2, 3, 4$, if $\varphi$ is unbounded it should be unbounded on positive elements. That is, there exist elements $a_n \in \mathcal{A}^+$ with $\|a_n\| = 1$ and $\varphi(a_n) > n$. Let $a = \sum_{k=1}^{\infty} \frac{a_k}{k^2} \in \mathcal{A}^+$. We have, since $a \geq \sum_{k=1}^{n} \frac{a_k}{k^2}$ and $\varphi$ is positive,

$$\varphi(a) \geq \varphi\Big(\sum_{k=1}^{n} \frac{a_k}{k^2}\Big) = \sum_{k=1}^{n} \frac{\varphi(a_k)}{k^2} \geq \sum_{k=1}^{n} \frac{1}{k}.$$

As this holds for all $n$, we have a contradiction. So $\varphi$ is bounded.

Because $\{e_j\}$ is positive and increasing and $|\varphi(e_j)| \leq \|\varphi\| \, \|e_j\| = \|\varphi\|$, we have that $\{\varphi(e_j)\}$ is bounded and non-decreasing, so $\lim_j \varphi(e_j)$ exists and $\lim_j \varphi(e_j) \leq \|\varphi\|$. Write $L = \lim_j \varphi(e_j)$. For any $a \in \mathcal{A}$ with $\|a\| \leq 1$, knowing that $\varphi$ is bounded,

$$|\varphi(a)|^2 = \lim_j |\varphi(e_j a)|^2 \leq \limsup_j \varphi(e_j^2)\varphi(a^*a) \leq \|\varphi\| \, \limsup_j \varphi(e_j) \leq L\, \|\varphi\|,$$

where we used that $e_j^2 \leq e_j$. This shows that $\|\varphi\| \leq L$, so $\|\varphi\| = L$.

(ii) $\Longrightarrow$ (i): Without loss of generality, we take $\|\varphi\| = 1$. We will first reduce to the unital case. Fix an approximate unit $\{e_j\}$. Let $\tilde{\varphi}$ be a Hahn–Banach extension of $\varphi$ to $\tilde{\mathcal{A}}$. Put $\alpha = \tilde{\varphi}(I_{\mathcal{A}})$. Then $|\alpha| \leq \|\tilde{\varphi}\| = \|\varphi\| = 1$. Since $\|2e_j - I_{\mathcal{A}}\| \leq 1$ (from $-I_{\mathcal{A}} \leq 2e_j - I_{\mathcal{A}} \leq I_{\mathcal{A}}$),

$$|2 - \alpha| = |2\|\varphi\| - \tilde{\varphi}(I_{\mathcal{A}})| = \lim_j |2\varphi(e_j) - \tilde{\varphi}(I_{\mathcal{A}})| = \lim_j |\tilde{\varphi}(2e_j - I_{\mathcal{A}})| \leq 1.$$

The inequalities $|\alpha| \leq 1$ and $|2-\alpha| \leq 1$, together, imply that $\alpha = 1$. So by replacing $\mathcal{A}$ with $\tilde{\mathcal{A}}$ we may assume that $\mathcal{A}$ is unital and $\varphi(I_{\mathcal{A}}) = 1$.

If $a \in \mathcal{A}^{\mathrm{sa}}$, and $\|a\| = 1$,

$$\begin{aligned}|\varphi(a) + in| = |\varphi(a + inI_{\mathcal{A}})| &\leq \|a + inI_{\mathcal{A}}\| = \|(a - inI_{\mathcal{A}})(a + inI_{\mathcal{A}})\|^{1/2} \\ &= \|a^2 + n^2\|^{1/2} = \operatorname{spr}\left((a^2 + n^2)^{1/2}\right) \\ &= (\operatorname{spr}(a)^2 + n^2)^{1/2} = (1 + n^2)^{1/2},\end{aligned}$$

for all $n \in \mathbb{Z}$. Looking at the imaginary part, $|\operatorname{Im}\varphi(a) + n| \leq (1+n^2)^{1/2}$. In particular, $\operatorname{Im}\varphi(a) \leq (1+n^2)^{1/2} - n$ for all $n$, so $\operatorname{Im}\varphi(a) \leq 0$. As the same argument works with $-a$ instead of $a$, $\operatorname{Im}\varphi(a) = 0$. So $\varphi$ is real. As $\|\varphi\| \leq 1$, $|\varphi(a)| \leq 1$; that is, $\varphi(a) \in [-1, 1]$.

If $a \in \mathcal{A}^+$ with $a \leq I_{\mathcal{A}}$, then $-I_{\mathcal{A}} \leq 2a - I_{\mathcal{A}} \leq I_{\mathcal{A}}$. Applying the above, $2\varphi(a) - 1 = \varphi(2a - I_{\mathcal{A}}) \in [-1, 1]$, so $\varphi(a) \in [0, 1]$, and $\varphi$ is positive. For arbitrary positive $a$, we use $a/\|a\|$. □

*11.5.5. Remark*. From the proof of Proposition 11.5.4, we also obtain $\|\varphi\| = \lim_j \varphi(e_j^2)$. Note that it is not clear that $\{e_j^2\}$ is an approximate unit, because squaring is not operator monotone and so the order might be lost. But

$$\|a - e_j^2 a\| \leq \|a - e_j a\| + \|e_j a - e_j^2 a\| \leq 2\|a - e_j a\| \to 0,$$

so the approximation still holds. And in the proof of Proposition 11.5.4 we can get the estimates, with $M = \limsup_j \varphi(e_j^2)$ and $m = \liminf_j \varphi(e_j)$,

$$|\varphi(a)^2| = \liminf_j |\varphi(e_j a) \leq \liminf_j \varphi(e_j^2)\varphi(a^* a) \leq m\,\|\varphi\|,$$

so $\|\varphi\| \leq m$. But $\varphi(e_j^2) \leq \|\varphi\|$, so $M \leq \|\varphi\|$. If follows that $\lim_j \varphi(e_j^2)$ exists and is equal to $\|\varphi\|$.

When $\mathcal{A}$ is not unital, states on the unitization do not need to restrict to states on $\mathcal{A}$. The canonical example of this is the state $\varphi(a, \lambda) = \lambda$, which restricts to zero on $\mathcal{A}$ (Exercise 11.5.12). It is true, on the other hand, that every state on $\mathcal{A}$ is the restriction of a state on $\tilde{\mathcal{A}}$. In fact,

***11.5.6. Proposition.***

*Let $\mathcal{A}$ be a non-unital $C^*$-algebra and $\tilde{\mathcal{A}}$ its unitization. For any $\varphi \in \mathcal{A}_+^*$ with $\|\varphi\| \leq 1$, there exists a unique extension $\tilde{\varphi} \in S(\tilde{\mathcal{A}})$ with $\tilde{\varphi}|_{\mathcal{A}} = \varphi$.*

*Proof.* Let $\varphi \in \mathcal{A}_+^*$ with $\|\varphi\| \leq 1$. On $\tilde{\mathcal{A}}$, define

$$\tilde{\varphi}(a, \lambda) = \varphi(a) + \lambda, \qquad a \in \mathcal{A}, \lambda \in \mathbb{C}.$$

Then $\tilde{\varphi}$ is obviously linear. We have, using the Kadison–Schwarz inequality (11.11),

$$\begin{aligned}\left|\tilde{\varphi}\big((a,\lambda)^*(a,\lambda)\big)\right| &= \varphi(a^*a) + 2\operatorname{Re}\bar{\lambda}\,\varphi(a) + |\lambda|^2 \\ &\geq |\varphi(a)|^2 + 2\operatorname{Re}\bar{\lambda}\,\varphi(a) + |\lambda|^2 \\ &= |\varphi(a)+\lambda|^2 \geq 0.\end{aligned}$$

Thus $\tilde{\varphi}$ is also positive, and by Proposition 11.5.4 it is a state since it is unital.

As for the uniqueness, the equality $\tilde{\varphi}(a,\lambda) = \tilde{\varphi}\big((a,0)+(0,\lambda)\big) = \varphi(a)+\lambda$ guarantees that only a single state extension is possible. □

Another strong consequence of Proposition 11.5.4 is the following. It looks rather innocuous, but it implies that states can be extended as states when there is a common unit, and also that when we consider a state $\varphi$ and an element $a \in \mathcal{A}$, we may think of $\varphi$ as a state on $C^*(a)$.

***11.5.7. Theorem.***

*Let $\mathcal{B} \subset \mathcal{A}$. If $\varphi \in S(\mathcal{B})$, then there exists $\tilde{\varphi} \in S(\mathcal{A})$ such that $\tilde{\varphi}|_{\mathcal{B}} = \varphi$.*

*Proof.* By considering $\mathcal{A}$ in its unitization $\tilde{\mathcal{A}}$, we have $\mathcal{B} \subset \tilde{\mathcal{A}}$. We obtain $\tilde{\varphi} \in \tilde{\mathcal{A}}^*$ by Hahn–Banach (Corollary 5.7.6), with $\tilde{\varphi}|_{\mathcal{B}} = \varphi$ and $\|\tilde{\varphi}\| = \|\varphi\| = 1$. Fix an approximate unit $\{e_j\}$ for $\mathcal{B}$. From Proposition 11.5.4 we have that

$$\|\tilde{\varphi}\| = \|\varphi\| = \lim_j \varphi(e_j) = \lim_j \tilde{\varphi}(e_j).$$

The inequalities $0 \leq e_j \leq I_{\tilde{\mathcal{A}}}$ give $\tilde{\varphi}(I_{\tilde{\mathcal{A}}}) \geq \tilde{\varphi}(e_j)$. In the limit we get $1 \geq \tilde{\varphi}(I_{\tilde{\mathcal{A}}}) \geq \|\varphi\| = 1$, so $\tilde{\varphi}(I_{\tilde{\mathcal{A}}}) = 1 = \|\tilde{\varphi}\|$. Then, using Proposition 11.5.4 again, we get that $\tilde{\varphi}$ is a state. Its restriction to $\mathcal{A}$ is still a state because it is positive and $\|\tilde{\varphi}|_{\mathcal{A}}\| = \|\varphi\| = 1$. □

The usefulness of Theorem 11.5.7, to be seen next, is that given $a$ normal we can construct a state on the abelian C$^*$-algebra $C^*(a)$ and extend it to a state on all of $\mathcal{A}$. We have now developed enough theory to show that states separate points. Indeed,

***11.5.8. Corollary.***

*Let $a \in \mathcal{A}$ be normal. Then there exists a state $\varphi \in S(\mathcal{A})$ such that $|\varphi(a)| = \|a\|$. In particular, states separate points.*

*Proof.* Assume first that $\mathcal{A}$ is unital. We know from Corollary 11.2.6 that we have a $*$-isomorphism $\rho : C^*(a) \to C(\sigma(a))$. By compactness, there exists $\lambda_0 \in \sigma(a)$ such that $|\lambda_0| = \max\{|\lambda| : \lambda \in \sigma(a)\} = \operatorname{spr}(a) = \|a\|$. Let $\tau_0 \in \Sigma(C(\sigma(a)))$ be given by $\tau(f) = f(\lambda_0)$, and $\tau : C^*(a) \to \mathbb{C}$ be given by $\tau(x) = \tau_0(\rho(x))$. Since both $\rho$ and $\tau_0$ are $*$-homomorphisms, so is $\tau$. And

$$|\tau(a)| = |\tau_0(\rho(a))| = |\tau_0(\mathrm{id}_{C(\sigma(a))})| = |\mathrm{id}_{C(\sigma(a))}(\lambda_0)| = |\lambda_0| = \|a\|.$$

So $\tau$ is a character. By Theorem 11.5.7, the state $\tau$ extends to $\varphi \in S(\mathcal{A})$ with $|\varphi(a)| = |\tau(a)| = \|a\|$. When $\mathcal{A}$ is not unital, we consider the restriction to $\mathcal{A}$ of the state obtained as above on the unitization $\tilde{\mathcal{A}}$. The equality $|\varphi(a)| = \|a\|$ guarantees that the restriction still satisfies $\|\varphi\| = 1$.

Now if $\varphi(a) = 0$, since $\varphi$ is positive we have that $0 = \varphi(\operatorname{Re} a) + i\varphi(\operatorname{Im} a)$ with $\varphi(\operatorname{Re} a), \varphi(\operatorname{Im} a) \in \mathbb{R}$. Then $0 = \varphi(\operatorname{Re} a) = \varphi(\operatorname{Im} a)$ and then $\operatorname{Re} a = \operatorname{Im} a = 0$ by the previous part of the proof, so $a = 0$. □

*11.5.9. Remark* . We cannot expect to have a state with $|\varphi(a)| = \|a\|$ for arbitrary $a \in \mathcal{A}$. For instance consider $\mathcal{A} = M_2(\mathbb{C})$ and $a = E_{12}$. The states of $\mathcal{A}$ are precisely the functionals $\varphi(a) = \operatorname{Tr}(ab)$ for some $b \geq 0$ with $\operatorname{Tr}(b) = 1$ (this follows, for instance, from Exercise 10.7.4). Then

$$\varphi(a) = b_{21}.$$

As $b \geq 0$ and $\operatorname{Tr}(b) = 1$, we have that $b_{11} + b_{22} = 1$, both non-negative, and $|b_{21}|^2 \leq b_{11}b_{22}$. Since $0 \leq b_{11}b_{22} = b_{11}(1 - b_{11}) \leq \frac{1}{4}$, it follows that $|b_{21}| \leq \frac{1}{2}$, and so it is impossible to have $\|a\| = 1 = \varphi(a)$. On the other hand, states do determine the norm of an arbitrary element (Exercise 11.5.6).

As we said above, Corollary 11.5.8 implies that states separate points; we mention in passing that states can also be used to characterize positivity:

***11.5.10. Corollary.***

*Let $\mathcal{A}$ be a $C^*$-algebra and $a \in \mathcal{A}$. The following statements are equivalent:*

*(i)* $a \geq 0$*;*

*(ii)* $\varphi(a) \geq 0$ *for all* $\varphi \in S(\mathcal{A})$*.*

*Proof.* (i) $\Longrightarrow$ (ii) By definition of state.

(ii) $\Longrightarrow$ (i) If $a = b + ic$ then given any $\varphi \in S(\mathcal{A})$ we have $\varphi(a) = \varphi(b) + i\varphi(c)$ with $\varphi(b), \varphi(c) \in \mathbb{R}$ since $\varphi$ is positive. The hypothesis gives us that $\varphi(c) = 0$ for all $\varphi \in S(\mathcal{A})$, and so $c = 0$ (Corollary 11.5.8). Hence $a$ is selfadjoint. If $a$ is not positive, there exists $\lambda \in \sigma(a)$ with $\lambda < 0$. The point evaluation $\delta_\lambda \in \Sigma(C_{\text{zero}}(\sigma(a))$ gives us—via Corollary 11.2.13—a state $\varphi_\lambda \in S(C^*(a))$ with $\varphi_\lambda(a) < 0$. Now we can extend $\varphi_\lambda$ to a state on $\mathcal{A}$ by Theorem 11.5.7, contradicting the hypothesis. Thus $a \geq 0$. □

**11.5.11. Definition.**

*A state $\varphi \in S(\mathcal{A})$ is **faithful** if $\varphi(a^*a) = 0$ implies $a = 0$.*

An easy and well-known example of a faithful state is the normalized trace on $M_n(\mathbb{C})$. Another very well-known example is $\varphi(f) = \int_0^1 f$ on $C[0,1]$.

***11.5.12. Proposition.***

*Let $\mathcal{A}$ be a separable $C^*$-algebra. Then $\mathcal{A}$ has a faithful state.*

*Proof.* Let $\{a_n'\} \subset \mathcal{A}$ be a countable dense subset of the closed unit ball of $\mathcal{A}$. Write $a_n' = a_n + ic_n$ with $a_n, c_n \in \mathcal{A}^{\text{sa}}$. The set $\{a_n\}$ is a countable dense subset of the unit ball of $\mathcal{A}^{\text{sa}}$; indeed, if $b = b^*$ and $\|b\| \leq 1$ there exists $n$ with $\|b - (a_n + ic_n)\| < \varepsilon$, and then

$$\|b - a_n\| = \|\text{Re}\,(b - (a_n + ic_n))\| \leq \|b - (a_n + ic_n)\| < \varepsilon.$$

By Corollary 11.5.8 there exist states $\varphi_n$ with $|\varphi_n(a_n)| = \|a_n\|$. Define

$$\varphi(a) = \sum_n \frac{1}{2^n}\,\varphi_n(a).$$

The series converges uniformly. Also, $\varphi$ is positive (limit of sums of positive) and for an approximate unit $\{e_j\}$ and using Proposition 11.5.4

$$\|\varphi\| = \lim_j \varphi(e_j) = \lim_j \sum_n \frac{1}{2^n}\,\varphi_n(e_j) = \sum_n \frac{1}{2^n}\,\lim_j \varphi_n(e_j) = \sum_n \frac{1}{2n} = 1$$

(the exchange of limits possible since the series converges uniformly), so $\varphi$ is a state. Now suppose that $a \in \mathcal{A}_+$, with $\varphi(a) = 0$; we may assume, without loss of generality, that $\|a\| \leq 1$. We have $\varphi_n(a) = 0$ for all $n$. Fix $\varepsilon > 0$; there exists $n$ with $\|a_n - a\| < \frac{\varepsilon}{2}$. Then

$$\|a\| \leq \|a - a_n\| + \|a_n\| < \frac{\varepsilon}{2} + \varphi(a_n) = \frac{\varepsilon}{2} + \varphi(a_n - a) \leq \frac{\varepsilon}{2} + \|a_n - a\| < \varepsilon.$$

As we can do this for any $\varepsilon > 0$, we get that $\|a\| = 0$, so $a = 0$ and $\varphi$ is faithful. □

*11.5.13. Example.* Here is a nice example of the calculation of the distance between two rank-one states. Let $\mathcal{H}$ be a separable Hilbert space, and $\xi, \eta \in \mathcal{H}$

with $\|\xi\| = \|\eta\| = 1$. Then

$$\phi(A) = \langle A\xi, \xi\rangle, \qquad \psi(A) = \langle A\eta, \eta\rangle$$

are states on $\mathcal{B}(\mathcal{H})$ (or any non-degenerate C*-algebra $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$, since we can extend states while preserving norm). We want to calculate $\|\phi - \psi\|$. Let $P, Q \in \mathcal{B}(\mathcal{H})$ be the orthogonal projections onto the span of $\xi, \eta$ respectively. We have

$$|\phi(A) - \psi(A)| = |\operatorname{Tr}(A(P-Q))| \leq \|A\| \, \|P-Q\|_1.$$

We get equality when $A = V^*$, where $P - Q = V|P-Q|$ is the polar decomposition. Thus

$$\|\phi - \psi\| = \|P - Q\|_1.$$

Using Exercise 10.7.14,

$$\|\phi - \psi\| = \|P - Q\|_1 = 2\sqrt{1 - |\langle \xi, \eta\rangle|^2}.$$

For arbitrary $\mathcal{A}$, the argument above survives as an inequality, the problem being that $V^*$ may fail to be in $\mathcal{A}$. An approximation would be enough, but the inequality can be strict; for instance, if $\mathcal{A} = \mathbb{C}$, then both states are equal and $\|\phi - \psi\| = 0$ even if $2\sqrt{1 - |\langle \xi, \eta\rangle|^2} > 0$.

The following technical lemma will be needed soon.

***11.5.14. Lemma.***

*Let $\mathcal{A}$ be a unital C*-algebra and $\varphi, \psi \in \mathcal{A}^*_+$. The following statements are equivalent:*

*(i)* $\|\varphi - \psi\| = \|\varphi\| + \|\psi\|$;

*(ii) for each* $\varepsilon > 0$ *there exists* $z \in \mathcal{A}^+$ *with* $\|z\| \leq 1$, *and such that* $\varphi(I_{\mathcal{A}} - z) < \varepsilon$, $\psi(z) < \varepsilon$.

*Proof.* (i) $\Longrightarrow$ (ii) Fix $\varepsilon > 0$. By definition of the norm (and multiplying by a suitable scalar if needed) there exists $y \in \mathcal{A}$ with $\|y\| = 1$ and $\varphi(y) - \psi(y) \geq \|\varphi - \psi\| - 2\varepsilon$. As $\varphi - \psi$ is Hermitian we may replace $y$ with $\operatorname{Re} y$, so we may assume without loss of generality that $y = y^*$. Let $z = \frac{1}{2}\,(I_{\mathcal{A}} + y) \geq 0$. Then, using Proposition 11.5.4,

$$\begin{aligned}
2\varphi(I_{\mathcal{A}} - z) + 2\psi(z) &= \varphi(I_{\mathcal{A}} - y) + \psi(I_{\mathcal{A}} + y) \\
&= \varphi(I_{\mathcal{A}}) + \psi(I_{\mathcal{A}}) - (\varphi(y) - \psi(y)) \\
&= \|\varphi\| + \|\psi\| - (\varphi(y) - \psi(y)) \\
&= \|\varphi - \psi\| - \big(\varphi(y) - \psi(y)\big)
\end{aligned}$$

$$\leq 2\varepsilon.$$

Since $I_{\mathcal{A}} - z \geq 0$ and $z \geq 0$, it follows from the positivity of $\varphi$ and $\psi$ that $\varphi(I_{\mathcal{A}} - z) < \varepsilon$, $\psi(z) < \varepsilon$.

(ii) $\Longrightarrow$ (i) Fix $\varepsilon > 0$ and $z$ as in the statement. Then, using that $\|2z - I_{\mathcal{A}}\| \leq 1$,

$$\begin{aligned}\|\varphi - \psi\| &\leq \|\varphi\| + \|\psi\| = \varphi(I_{\mathcal{A}}) + \psi(I_{\mathcal{A}})\\ &\leq \varphi(I_{\mathcal{A}}) + \psi(I_{\mathcal{A}}) + 2\varepsilon - 2\varphi(I_{\mathcal{A}} - z) + 2\varepsilon - 2\psi(z)\\ &= (\varphi - \psi)(2z - I_{\mathcal{A}}) + 4\varepsilon \leq \|\varphi - \psi\| + 4\varepsilon.\end{aligned}$$

As $\varepsilon$ was arbitrary, we get that $\|\varphi - \psi\| = \|\varphi\| + \|\psi\|$. □

As mentioned in [Ped79, 3.2.4], in the non-unital case Lemma 11.5.14 still holds but the condition (ii) has to be replaced by the existence, for every $\varepsilon > 0$, of positive $x, y \in \mathcal{A}$ with $x + y \leq I_{\mathcal{A}}$ and $\varphi(x) \geq \|\varphi\| - \varepsilon$, $\psi(y) \geq \|\psi\| - \varepsilon$ (Exercise 11.5.8).

***11.5.15. Proposition.***

*Let $\mathcal{A}$ be a unital C*-algebra. Then $S(\mathcal{A})$ is weak*-compact.*

*Proof.* We know that the unit ball of $\mathcal{A}^*$ is weak*-compact by Banach–Alaoglu (Theorem 7.2.13). So it is enough to show that $S(\mathcal{A})$ is weak*-closed. Let $\{\psi_j\}$ be a net in $S(\mathcal{A})$ such that $\psi_j \xrightarrow{\text{weak}^*} \psi$. We want to show that $\psi \in S(\mathcal{A})$. If $a \in \mathcal{A}_+$,

$$\psi(a) = \lim_j \psi_j(a) \geq 0,$$

so $\psi \geq 0$. And $\psi(I_{\mathcal{A}}) = \lim_j \psi(I_{\mathcal{A}}) = 1$, so $\psi \in S(\mathcal{A})$. □

The non-unital case will be discussed in Proposition 11.5.26 and beyond.

We saw on page 215 how to decompose a signed measure as a difference of two positive measures. We also saw in Proposition 11.3.10 that we can write a selfadjoint element in a C*-algebra as a difference of two positive elements. It should come as no big surprise that a similar decomposition is available in the dual of a C*-algebra. There is also a polar decomposition for linear functionals, but it only makes sense in the context of normal functionals on a von Neumann algebra (Proposition 12.6.4).

**11.5.16. Proposition.** *(Jordan Decomposition of a real functional)*

*Let $\mathcal{A}$ be a $C^*$-algebra, and $\phi \in \mathcal{A}^*$ Hermitian. Then there exist unique $\phi_1, \phi_2 \in \mathcal{A}^*$, positive, such that*

$$\phi = \phi_1 - \phi_2, \qquad \|\phi\| = \|\phi_1\| + \|\phi_2\|.$$

*Proof.* Assume first that $\mathcal{A}$ is unital. It is clear that $\overline{\mathrm{conv}}^{w^*}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big]$ consists of Hermitian functionals of norm at most one. Suppose that $\psi \in \mathcal{A}^*$ is Hermitian, $\|\psi\| \le 1$, and

$$\psi \in \mathcal{A}^* \setminus \overline{\mathrm{conv}}^{w^*}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big]$$

By Hahn–Banach (Theorem 5.7.18) and Proposition 7.2.10, there exist $a \in \mathcal{A}$ and $d \in \mathbb{R}$ such that

$$\operatorname{Re}\psi(a) < d < \operatorname{Re}\varphi(a), \qquad \varphi \in \overline{\mathrm{conv}}^{w^*}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big].$$

Using that $\psi$ is Hermitian, and replacing $a$ with its real part, we can write

$$\psi(a) < d < \varphi(a), \qquad \varphi \in \overline{\mathrm{conv}}^{w^*}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big].$$

In particular $\psi(a) < d < -\varphi(a), \qquad \varphi \in S(\mathcal{A})$, so

$$\varphi(a) < -d < -\psi(a), \qquad \varphi \in S(\mathcal{A}).$$

But then, taking $\varphi$ to be the one from Corollary 11.5.8,

$$\|a\| = \varphi(a) < -d < -\psi(a) \le \|-\psi\|\,\|a\| = \|a\|,$$

a contradiction. It follows that

$$\{\psi \in \mathcal{A}^*, \quad \text{Hermitian}, \ \|\psi\| \le 1\} = \overline{\mathrm{conv}}^{w^*}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big]. \tag{11.13}$$

Our assumption that $\mathcal{A}$ is unital lets us use that $S(\mathcal{A})$ is compact (Proposition 11.5.15); then so is $\mathrm{conv}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big]$, by Lemma 7.5.14. Hence

$$\{\psi \in \mathcal{A}^*, \quad \text{Hermitian}, \ \|\psi\| \le 1\} = \mathrm{conv}\big[S(\mathcal{A}) \cup [-S(\mathcal{A})]\big]. \tag{11.14}$$

Consider $\phi \in \mathcal{A}^*$, Hermitian. By (11.13) there exist $\varphi_1, \varphi_2 \in S(\mathcal{A})$ and $\alpha, \beta \ge 0$, with $\alpha + \beta = 1$ and

$$\frac{\phi}{\|\phi\|} = \alpha\varphi_1 - \beta\varphi_2.$$

Then, using that $\varphi_1, \varphi_2 \in S(\mathcal{A})$,

$$\left\|\frac{\phi}{\|\phi\|}\right\| = 1 = \alpha + \beta = \|\alpha\varphi_1\| + \|\beta\varphi_2\|.$$

If we now take $\phi_1 = \alpha\|\phi\|\,\varphi_1$, $\phi_2 = \beta\|\phi\|\varphi_2$, we have $\phi = \phi_1 - \phi_2$, and

$$\|\phi\| = \alpha\|\phi\| + \beta\|\phi\| = \|\phi_1\| + \|\phi_2\|.$$

Now the case where $\mathcal{A}$ is non-unital. Let $\phi \in \mathcal{A}^*$, Hermitian, $\|\phi\| \leq 1$. We extend by Hahn–Banach (Corollary 5.7.6) to a linear functional $\phi'$ on the unitization $\tilde{\mathcal{A}}$, with $\|\phi'\| = \|\phi\|$. It might not be the case that $\phi'$ is Hermitian, but we can define

$$\tilde{\phi}(a) = \tfrac{1}{2}\big(\phi'(a) + \overline{\phi'(a^*)}\big).$$

By construction, $\tilde{\phi}$ is Hermitian. Also, $\|\tilde{\phi}(a)\| \leq \|\phi'(a)\|$. But it agrees with $\phi$ on $\mathcal{A}$, so $\|\tilde{\phi}\| = \|\phi'\| = \|\phi\|$. Because we are on the unitization, the first part of the proof implies that there exist $\phi_1', \phi_2' \in \tilde{\mathcal{A}}^*$, positive, such that $\tilde{\phi} = \phi_1' - \phi_2'$ and $\|\tilde{\phi}\| = \|\phi_1'\| + \|\phi_2'\|$. If we denote by $\phi_1, \phi_2$ their restrictions to $\mathcal{A}$, we get

$$\phi(a) = \phi_1(a) - \phi_2(a), \qquad a \in \mathcal{A}.$$

The problem is that a priori we do not know that $\phi_1$ and $\phi_2$ retain the norm when restricted to $\mathcal{A}$; their norms could be smaller (see the paragraph before Proposition 11.5.6). But

$$\|\phi_1\| + \|\phi_2\| \leq \|\phi_1'\| + \|\phi_2'\| = \|\tilde{\phi}\| = \|\phi\| \leq \|\phi_1\| + \|\phi_2\|.$$

This forces $\|\phi\| = \|\phi_1\| + \|\phi_2\|$, as needed.

Now it remains to deal with the uniqueness. Suppose that $\phi_1, \phi_2, \psi_1, \psi_2$ are positive linear functionals with $\phi_1 - \phi_2 = \psi_1 - \psi_2$ and $\|\phi_1\| + \|\phi_2\| = \|\phi_1 - \phi_2\| = \|\psi_1\| + \|\psi_2\|$. If we extend these functionals to the unitization of $\mathcal{A}$ as in Proposition 11.5.6, the equality $\phi_1 + \psi_2 = \psi_1 + \phi_2$ will still hold in the unitization, and the norms are unchanged; all in all, this shows that we may assume without loss of generality that $\mathcal{A}$ is unital. Fix $\varepsilon > 0$ and choose $z$ as in Lemma 11.5.14, so that $0 \leq z \leq I_{\mathcal{A}}$ and $\phi_1(I_{\mathcal{A}} - z) < \varepsilon$, $\phi_2(z) < \varepsilon$. We have

$$\psi_1(z) \geq \psi_1(z) - \psi_2(z) = \phi_1(z) - \phi_2(z) \geq \phi_1(I_{\mathcal{A}}) - 2\varepsilon,$$

and

$$\psi_2(I_{\mathcal{A}} - z) \geq \psi_2(I_{\mathcal{A}} - z) - \psi_1(I_{\mathcal{A}} - z) = \phi_2(I_{\mathcal{A}} - z) - \phi_1(I_{\mathcal{A}} - z) \geq \phi_2(I_{\mathcal{A}}) - 2\varepsilon.$$

Then, using Proposition 11.5.4,

$$\begin{aligned} \psi_1(I_{\mathcal{A}} - z) + \psi_2(z) &\leq \psi_1(I_{\mathcal{A}}) - \phi_1(I_{\mathcal{A}}) + \psi_2(I_{\mathcal{A}}) - \phi_2(I_{\mathcal{A}}) + 4\varepsilon \\ &= \|\psi_1\| + \|\psi_2\| - \|\phi_1\| - \|\phi_2\| + 4\varepsilon = 4\varepsilon. \end{aligned} \tag{11.15}$$

For any $a \in \mathcal{A}$, and using that $(I_{\mathcal{A}} - z)^2 \leq I_{\mathcal{A}} - z$,

$$\begin{aligned} |\phi_1(a(I_{\mathcal{A}} - z))|^2 &\leq \phi_1(aa^*)\phi_1((I_{\mathcal{A}} - z)^2) \leq \|a\|^2\, \|\phi_1\|\, \phi_1(I_{\mathcal{A}} - z) \\ &\leq \|a\|^2\, \|\phi_1 - \phi_2\|\, \phi_1(I_{\mathcal{A}} - z). \end{aligned}$$

With the same idea,

$$|\phi_2(az)|^2 \leq \|a\|^2\, \|\phi_1 - \phi_2\|\, \phi_2(z).$$

Also, from (11.15), and with $c = \max\{\|\psi_1\|^{1/2}, \|\psi_2\|^{1/2}\}$,

$$|\psi_1(a(I_{\mathcal{A}} - z)) + \psi_2(az)| \leq |\psi_1(a(I_{\mathcal{A}} - z))| + |\psi_2(az)|$$

$$
\begin{aligned}
&\leq \psi_1(aa^*)^{1/2}\psi_1((I_{\mathcal{A}}-z)^2)^{1/2}+\psi_2(aa^*)^{1/2}\psi_2(z^2)^{1/2}\\
&\leq \|a\|\,\|\psi_1\|^{1/2}\,\psi_1(I_{\mathcal{A}}-z)^{1/2}+\|a\|\,\|\psi_2\|^{1/2}\,\psi_2(z)^{1/2}\\
&\leq \|a\|\,c\,\bigl(\psi_1(I_{\mathcal{A}}-z)^{1/2}+\psi_2(z)^{1/2}\bigr)\\
&\leq 2\|a\|\,c\,\bigl(\psi_1(I_{\mathcal{A}}-z)+\psi_2(z)\bigr)\\
&\leq 8\|a\|\,c\,\varepsilon.
\end{aligned}
$$

Now

$$
\begin{aligned}
|\phi_1(a)-\psi_1(a)| &= \bigl|\phi_1(az)-\psi_1(az)+\phi_1(a(I_{\mathcal{A}}-z))-\psi_1(a(I_{\mathcal{A}}-z))\bigr|\\
&= \bigl|\phi_2(az)-\psi_2(az)+\phi_1(a(I_{\mathcal{A}}-z))-\psi_1(a(I_{\mathcal{A}}-z))\bigr|\\
&= \bigl|\phi_1(a(I_{\mathcal{A}}-z))+\phi_2(az)-\psi_1(a(I_{\mathcal{A}}-z)-\psi_2(az))\bigr|\\
&\leq 2\|a\|\,\|\phi_1-\phi_2\|^{1/2}\,\varepsilon^{1/2}+8\|a\|\,c\,\varepsilon.
\end{aligned}
$$

As this can be done for any $\varepsilon>0$, it follows that $\phi_1=\psi_1$; and this forces $\phi_2=\psi_2$. ☐

The argument at the end of the proof of existence in Proposition 11.5.16 also implies that $\|\phi_1\|=\|\phi_1'\|$, $\|\phi_2\|=\|\phi_2'\|$, which is interesting in light of the already mentioned fact that states in the unitization $\tilde{\mathcal{A}}$ do not necessarily restrict to states on $\mathcal{A}$.

***11.5.17. Corollary.***

*Let $\varphi\in\mathcal{A}^*$. Then there exist $\alpha_j\in\{0\}\cup\mathbb{R}^+$, $\varphi_j\in S(\mathcal{A})$, $j=1,\ldots,4$, such that $\varphi=\alpha_1\varphi_1-\alpha_2\varphi_2+i(\alpha_3\varphi_3-\alpha_4\varphi_4)$.*

*Proof.* Let

$$
\varphi_R(x)=\frac{\varphi(x)+\overline{\varphi(x^*)}}{2},\qquad \varphi_I(x)=\frac{\varphi(x)-\overline{\varphi(x^*)}}{2i}.
$$

It is straightforward to check that $\varphi_R$ and $\varphi_I$ are Hermitian (Exercise 11.5.2). By Proposition 11.5.16, there exist $\varphi_1',\varphi_2'\in\mathcal{A}^+$ with $\varphi_R=\varphi_1'-\varphi_2'$, and $\varphi_3',\varphi_4'\in\mathcal{A}^+$ with $\varphi_I=\varphi_3'-\varphi_4'$. Then, putting $\varphi_j=\varphi_j'/\|\varphi_j'\|\in S(\mathcal{A})$,

$$
\varphi=\varphi_R+i\varphi_I=\|\varphi_1'\|\,\varphi_1-\|\varphi_2'\|\,\varphi_2+i(\|\varphi_3'\|\,\varphi_3-\|\varphi_4'\|\,\varphi_4).
$$

☐

*11.5.18. Remark* . Proposition 11.5.16 (and thus Corollary 11.5.17) can be proven without that many technicalities if one drops the condition on the norms—but we will need it soon. We outline the argument in Exercise 11.5.7.

**11.5.19. Definition.**

*Let $\mathcal{A}$ be a $C^*$-algebra and $\mathcal{B} \subset \mathcal{A}$ a $C^*$-subalgebra. We say that $\mathcal{B}$ is* ***hereditary*** *if $0 \leq a \leq b$ with $b \in \mathcal{B}$, implies $a \in \mathcal{B}$.*

An example of hereditary subalgebra is given by ideals. Indeed, if $\mathcal{J} \subset \mathcal{A}$ is an ideal and $0 \leq a \leq b$ with $b \in \mathcal{J}$, we may consider the quotient map $\rho : \mathcal{A} \to \mathcal{A}/\mathcal{J}$. This is a $*$-homomorphism, and we have $0 \leq \rho(a) \leq \rho(b) = 0$, so $\rho(a) = 0$ and then $a \in \ker \rho = \mathcal{J}$. Hereditary subalgebras which are not ideals do exist, though. For instance if $p \in \mathcal{A}$ is a projection, then $p\mathcal{A}p$ is a hereditary subalgebra of $\mathcal{A}$ (Exercise 11.5.9). More generally, for any $a \in \mathcal{A}^+$ the subalgebra $\overline{a\mathcal{A}a}$ is hereditary (Exercise 11.5.10); it may not be a proper subalgebra, though, see Proposition 11.5.23.

The following lemma contains a technical fact about hereditary subalgebras.

***11.5.20. Lemma.***

*Let $\mathcal{A}$ be a unital $C^*$-algebra and $\mathcal{B} \subset \mathcal{A}$ a hereditary $C^*$-subalgebra. Let $\psi \in S(\mathcal{A})$ and $\varphi \in \mathcal{B}^*$ given by $\varphi = \psi|_{\mathcal{B}}$, and $\{e_j\}$ an approximate unit in $\mathcal{B}$. Then*

$$\psi(a) = \lim_j \psi(e_j a e_j) = \lim_j \varphi(e_j a e_j), \qquad a \in \mathcal{A}. \tag{11.16}$$

*Proof.* Given an approximate unit $\{e_j\}$ in $\mathcal{B}$, we have $e_j \leq I_{\mathcal{A}}$ for all $j$, so

$$1 \geq \psi(I_{\mathcal{A}}) \geq \limsup_j \psi(e_j) = \lim_j \varphi(e_j) = 1.$$

That is,

$$1 = \psi(I_{\mathcal{A}}) = \lim_j \varphi(e_j) = \lim_j \varphi(e_j^2).$$

As $(I_{\mathcal{A}} - e_j)^2 \leq I_{\mathcal{A}} - e_j$,

$$\begin{aligned}
|\psi(a - e_j a e_j)| &\leq |\psi(a - e_j a)| + |\psi(e_j a - e_j a e_j)| \\
&= |\psi((I_{\mathcal{A}} - e_j)a)| + |\psi(e_j a (I_{\mathcal{A}} - e_j)) \\
&\leq \psi(aa^*)^{1/2} \psi(I_{\mathcal{A}} - e_j)^{1/2} + \psi(e_j a a^* e_j)^{1/2} \psi(I_{\mathcal{A}} - e_j)^{1/2} \\
&\leq 2\|a\| \, \|\psi\|^{1/2} \psi(I_{\mathcal{A}} - e_j)^{1/2} \to 0
\end{aligned}$$

If $a \in \mathcal{A}^+$, then $0 \leq e_j a e_j \leq \|a\| \, e_j^2 \in \mathcal{B}^+$. As $\mathcal{B}$ is hereditary, it follows that $e_j a e_j \in \mathcal{B}$ for all $a \in \mathcal{A}^+$, and since every element is a linear combination of positives, $e_j a e_j \in \mathcal{B}$ for all $a \in \mathcal{A}$ and all $j$. Thus

$$\psi(a) = \lim_j \psi(e_j a e_j) = \lim_j \varphi(e_j a e_j), \qquad a \in \mathcal{A},$$

where $e_j a e_j \in \mathcal{B}$ for all $a \in \mathcal{A}$. □

Now we can prove that states extend uniquely from hereditary subalgebras:

***11.5.21. Proposition.***

*Let $\mathcal{A}$ be a $C^*$-algebra and $\mathcal{B} \subset \mathcal{A}$ a hereditary $C^*$-subalgebra. Let $\varphi \in S(\mathcal{B})$. Then there exists a unique $\psi \in S(\mathcal{A})$ such that $\psi|_{\mathcal{B}} = \varphi$*

*Proof.* If $\mathcal{A}$ is not unital, we replace $\mathcal{A}$ with $\tilde{\mathcal{A}}$. The algebra $\mathcal{B}$ is still hereditary in $\tilde{\mathcal{A}}$ (Exercise 11.5.13), so we may assume without loss of generality that $\mathcal{A}$ is unital. The existence of the extension can then be achieved by extending $\varphi$ by Hahn–Banach (Corollary 5.7.6) to obtain $\psi \in \mathcal{A}^*$ with $\|\psi\| = \|\varphi\| = 1$. By Lemma 11.5.20, $\psi(I_{\mathcal{A}}) = 1$ and so $\psi \geq 0$ (and hence a state) by Proposition 11.5.4.

The nontrivial fact here is the uniqueness of the extension, but we have already done the computations. Indeed, suppose that $\psi \in S(\mathcal{A})$ with $\psi|_{\mathcal{B}} = \varphi$. By Lemma 11.5.20

$$\psi(a) = \lim_j \psi(e_j a e_j) = \lim_j \varphi(e_j a e_j), \qquad a \in \mathcal{A}.$$

So $\psi$ is uniquely determined by $\varphi$. □

As simple as the proof of Proposition 11.5.15 is, it cannot be carried in the non-unital case, as we will see. First we need some material.

**11.5.22. Definition.**

*An element $a \in \mathcal{A}^+$ is **strictly positive** if $\varphi(a) > 0$ for all $\varphi \in S(\mathcal{A})$.*

Strict positivity is meaningful only for non-unital algebras, in the sense that “strictly positive” is equivalent to “positive and invertible” in the unital case. The next two results contain all the information we will need about strictly positive elements.

***11.5.23. Proposition.***

*Let $a \in \mathcal{A}^+$. The following statements are equivalent:*

*(i) $a$ is strictly positive;*

*(ii) $\overline{a\mathcal{A}a} = \mathcal{A}$;*

*(iii) if $\mathcal{A}$ is unital, $a$ is invertible.*

*Proof.* (i) $\implies$ (ii) Suppose that $\overline{a\mathcal{A}a} \subsetneq \mathcal{A}$. As $\mathcal{A}$ is spanned by its positive elements, there exists nonzero $b \in \mathcal{A}^+ \setminus \overline{a\mathcal{A}a}$. By Hahn–Banach (Corollary 5.7.19) there exists $\varphi \in \mathcal{A}^*$ such that

$$\varphi(b) = 1 \qquad \text{and} \qquad \varphi(aca) = 0 \quad \text{for all } c \in \mathcal{A}. \tag{11.17}$$

Since $\operatorname{Re}\varphi$ also satisfies (11.17), we may assume without loss of generality that $\varphi$ is Hermitian. Using Proposition 11.5.16 we can write $\varphi = \alpha_1\psi_1 - \alpha_2\psi_2$ with $\alpha_1, \alpha_2 \geq 0$, $\psi_1, \psi_2 \in S(\mathcal{A})$, and $\|\alpha_1\psi_1\| + \|\alpha_2\psi_2\| = \|\varphi\|$. So $\alpha_1\psi_1(b) > 0$ (which implies $\psi_1(b) > 0$) and $\alpha_1\psi_1(aca) = \alpha_2\psi_2(aca)$ for all $c \in \mathcal{A}$.

Fix an approximate unit $\{e_j\}$ of $\overline{a\mathcal{A}a}$. By Proposition 11.5.4 and Remark 11.5.5,

$$\|\psi_1|_{\overline{a\mathcal{A}a}}\| = \lim_j \psi_1(e_j) = \lim_j \psi_1(e_j^2).$$

Fix $\varepsilon > 0$. By Lemma 11.5.14 there exists $z \in \tilde{\mathcal{A}}$ such that, still writing $\psi_1, \psi_2$ for their extensions to the unitization $\tilde{\mathcal{A}}$ (which preserve the norm by Proposition 11.5.6), $\alpha_1\psi_1(I_{\mathcal{A}} - z) < \varepsilon$ and $\alpha_2\psi_2(z) < \varepsilon$. By Lemma 11.5.20,

$$\psi_2(z) = \lim_j \psi_2(e_j z e_j) \qquad\qquad \psi_1(I_{\mathcal{A}} - z) = \lim_j \psi_1(e_j(I_{\mathcal{A}} - z)e_j).$$

Then, for $j$ big enough,

$$\alpha_1\psi_1(e_j z e_j) = \alpha_2\psi_2(e_j z e_j) < \varepsilon,$$

and so (still for $j$ big enough)

$$\varepsilon > \alpha_1\psi_1(e_j(I_{\mathcal{A}} - z)e_j) = \alpha_1\psi_1(e_j^2) - \alpha_1\psi_1(e_j z e_j) > \alpha_1\psi_1(e_j^2) - \varepsilon.$$

This shows that $0 = \lim_j \psi_1(e_j^2) = \|\psi_1|_{\overline{a\mathcal{A}a}}\|$. That is, $\psi_1|_{\overline{a\mathcal{A}a}} = 0$. In particular, $\psi_1(a^3) = 0$. Then

$$0 \leq \psi_1(a^4) = \psi_1(a^{3/2} a\, a^{3/2}) \leq \|a\|\, \psi_1(a^3) = 0.$$

Now, using the Kadison–Schwarz inequality (11.11) twice,

$$0 \leq \psi_1(a)^2 \leq \psi_1(a^2) \leq \psi_1(a^4)^{1/2} = 0.$$

So $a$ is not strictly positive.

(ii) $\implies$ (i) Suppose that $\varphi(a) = 0$ for some $\varphi \in S(\mathcal{A})$. Then, for any $b \in \mathcal{A}$, using Cauchy–Schwarz we have

$$|\varphi(aba)|^2 \leq \varphi(ab^*ba)\phi(a^2) \leq \|a\|\, \varphi(b^*a^2b)\phi(a) = 0.$$

As elements of the form $aba$ are dense in $\mathcal{A}$, we conclude that $\varphi = 0$, a contradiction. This shows that $\varphi(a) > 0$ for any state $\varphi$.

(iii) If $\mathcal{A}$ is unital and $a$ is positive and invertible, then $a \geq \varepsilon I_{\mathcal{A}}$, where $\varepsilon = \min\sigma(a) > 0$. Then, for any $\varphi \in S(\mathcal{A})$ we have $\varphi(a) \geq \varepsilon\varphi(I_{\mathcal{A}}) = \varepsilon > 0$. Conversely, if $a$ is strictly positive then $\overline{a\mathcal{A}a} = \mathcal{A}$. So there exists $b \in \mathcal{A}$ with $\|I_{\mathcal{A}} - aba\| < 1$. By Proposition 6.2.3 we get that $aba$ is invertible. So there exists $c \in \mathcal{A}$ with $caba = I_{\mathcal{A}}$ and $abac = I_{\mathcal{A}}$. Thus $cab$ is a left inverse and $bac$ is a right inverse for $a$; this implies that $cab = bac$ and that $a$ is invertible. $\square$

***11.5.24. Proposition.***

*Let $\mathcal{A}$ be a $C^*$-algebra. The following statements are equivalent.*

*(i) $\mathcal{A}$ contains a countable approximate unit;*

*(ii) $\mathcal{A}$ has a strictly positive element.*

*Proof.* (i) $\implies$ (ii) Let $\{e_n\}$ be a countable approximate unit for $\mathcal{A}$, and let $a = \sum_n 2^{-n} e_n$. Suppose that $\varphi(a) = 0$ for some $\varphi \in S(\mathcal{A})$. Then for any $n \in \mathbb{N}$ we have $0 \leq \varphi(e_n) \leq 2^n \varphi(a) = 0$, so $\varphi(e_n) = 0$. By Proposition 11.5.4, for $b \in \mathcal{A}^+$

$$\begin{aligned}\phi(b) = \lim_n \phi(e_n b e_n) &= \lim_n \phi(e_n^{1/2}[e_n^{1/2} b e_n^{1/2}]e_n^{1/2}) \\ &\leq \limsup_n \|e_n^{1/2} b e_n^{1/2}\| \, \phi(e_n) \leq \|b\| \, \limsup_n \phi(e_n) = 0.\end{aligned}$$

Thus $\varphi = 0$ on $\mathcal{A}^+$, and so $\varphi = 0$. In other words, $\varphi(a) > 0$ for all $\varphi \in S(\mathcal{A})$.

(ii) $\implies$ (i) Let $a \in \mathcal{A}^+$ be strictly positive. Working on the unitization $\tilde{\mathcal{A}}$, we can form $e_n = a(a + \frac{1}{n} I_{\mathcal{A}})^{-1}$. Note that $0 \leq e_n \leq e_{n+1} \leq I_{\mathcal{A}}$ by Lemma 11.4.3. Because $\mathcal{A}$ is an ideal in $\tilde{\mathcal{A}}$, we have $e_n \in \mathcal{A}$ for all $n$. Let $f_n : \sigma(a) \to [0, \infty)$ be given by $f_n(t) = t^2/(t + 1/n)$. Then $\lim_n f_n(t) = t$ for all $t$. By Dini's Theorem (Theorem 2.6.15), the convergence is uniform. Writing $f$ for the identity function $f(t) = t$,

$$\|a - a e_n\| = \|f(a) - f_n(a)\| \leq \|f - f_n\|_\infty \to 0.$$

Now, given $b \in \mathcal{A}$, since $a$ is strictly positive there exists a sequence $\{b_n\} \subset \mathcal{A}$ with $a b_n a \to b$. Then

$$\begin{aligned}\|b - b e_k\| &\leq \|b - a b_n a\| + \|a b_n a - a b_n a e_k\| + \|a b_n a e_k - b e_k\| \\ &\leq \|b - a b_n a\| + \|a b_n a - a b_n a e_k\| + \|a b_n a - b\| \|e_k\| \\ &\leq 2\|b - a b_n a\| + \|a b_n a - a b_n a e_k\|.\end{aligned}$$

Then

$$\limsup_k \|b - b e_k\| \leq 2\|b - a b_n a\|.$$

As this occurs for all $n$, by the Limsup Routine $\lim_k \|b - b e_k\| = 0$, so $\{e_k\}$ is an approximate unit for $\mathcal{A}$. ☐

***11.5.25. Corollary.***

*Let $\mathcal{A}$ be a $C^*$-algebra, $a \in \mathcal{A}$ strictly positive, and $p$ a projection. Then $pap$ is invertible in $p\mathcal{A}p$.*

*Proof.* By Exercise 11.5.15, the sequence $\{a^{1/n}\}_n$ is an approximate unit for $\mathcal{A}$. Hence $pa^{1/n}p \to p$, which by Proposition 6.2.3 implies that $pa^{1/n}p$ is invertible in $p\mathcal{A}p$ for all big enough $n$. In particular there exists $k \in \mathbb{N}$ such that $pa^{2^{-k}}p$ is invertible in $p\mathcal{A}p$. As

$$(pa^{2^{-k}}p)^2 = pa^{2^{-k}}pa^{2^{-k}}p \le pa^{2^{-k+1}}p,$$

it follows that $pa^{2^{-k+1}}p$ is invertible. Inductively, $pap = pa^{2^{-k+k}}p$ is invertible in $p\mathcal{A}p$. $\square$

The machinery of strictly positive elements allows us to prove this striking counterpart to Proposition 11.5.15.

**11.5.26. Proposition.**

*Let $\mathcal{A}$ be a non-unital $C^*$-algebra. Then*

$$\overline{S(\mathcal{A})}^{w^*} = \{\varphi \in \mathcal{A}^* : \ \varphi \ge 0, \ \|\varphi\| \le 1\}.$$

*In particular,* $0 \in \overline{S(\mathcal{A})}^{w^*}$.

*Proof.* Suppose that $0 \notin \overline{S(\mathcal{A})}^{w^*}$. We consider the locally convex space $\mathcal{A}^*$ with its weak*-topology. By Proposition 7.2.10, its dual is $\mathcal{A}$. Using the second geometric form of the Hahn–Banach (Theorem 5.7.18) on the compact convex set $\{0\}$ and the closed—and also compact, but we do not need this here—convex set $\overline{S(\mathcal{A})}^{w^*}$, there exists $a \in \mathcal{A}$ and $c > 0$ such that

$$0 < c < \operatorname{Re}\varphi(a), \qquad \varphi \in \overline{S(\mathcal{A})}^{w^*}.$$

Since each $\varphi$ is positive, we have $2\operatorname{Re}\varphi(a) = \varphi(a) + \varphi(a^*) = \varphi(a + a^*)$. This allows us to replace $a$ with $\operatorname{Re} a$; in other words, we may assume that $a$ is selfadjoint. So now we have $0 < c < \varphi(a)$ for all states $\varphi$, which implies that $a$ is positive (Corollary 11.5.10). This $a$ is strictly positive and $\varphi(a) \ge c$ for all $\varphi \in \overline{S(\mathcal{A})}^{w^*}$. Using the Gelfand transform (Corollary 11.2.13), we know that $C^*(a) \simeq C_{\text{zero}}(\sigma(a))$. This algebra cannot be unital, because $a$ is strictly positive. Indeed, if $p \in C^*(a)$ is the unit, then $pa = a$; using Proposition 11.5.23, for any $b \in \mathcal{A}$ we have $b = \lim ab_n a$, and then $pb = \lim pab_n a = \lim ab_n a = b$, and so—after an analog computation on the right—$\mathcal{A}$ would be unital, a contradiction.

Let us denote by $\rho : C^*(a) \to C_{\text{zero}}(\sigma(a))$ the $*$-isomorphism from Corollary 11.2.13. Knowing that $C_{\text{zero}}(\sigma(a))$ is not unital, we can construct a net $\{t_j\}$ as follows. Let $\mathcal{F}_0$ be the family of finite subsets of $C_{\text{zero}}(\sigma(a))$. Put $\mathcal{F} = \{(F, \delta) : \ F \in \mathcal{F}_0, \ \delta > 0\}$ with the order $(F, \delta) \le (F', \delta')$ if $F \subset F'$ and $\delta' \le \delta$. Given any $j = (F, \delta) \in \mathcal{F}$ there exists $K \subset T$, compact, such that $|f(t)| < \delta$ on $T \setminus K$ for all $f \in F$. Choose $t_j \in T \setminus K$. Then the net of point-evaluations $\varphi_j(f) = f(t_j)$ converges pointwise to 0; indeed, for fixed $f \in C(T)$ and $\varepsilon > 0$, let $j_0 = (\{f\}, \varepsilon)$. If $j = (F, \delta) \geq j_0$, this means that $f \in F$ and $\delta \leq \varepsilon$, giving us $|f(t_j)| < \delta \leq \varepsilon$. Hence $\varphi_j(f) = f(t_j) \to 0$ for all $f$; this means that $\varphi_j \circ \rho \xrightarrow{\text{weak}^*} 0$ and so $0 \in \overline{S(C_{\text{zero}}(\sigma(a)))}^{w^*}$. Using Theorem 11.5.7 we can extend each $\psi_j = \varphi_j \circ \rho$ to all of $\mathcal{A}$. Now,

$$|\psi_j(aba)| \leq \psi_j(abb^*a)^{1/2}\, \psi_j(a^2)^{1/2} \leq \|a\|\,\|b\|\,\psi_j(a^2)^{1/2},$$

since $\|psi_j\| = 1$ for all $j$. As $\psi_j(a^2) \to 0$, we get that $\psi_j(aba) \to 0$. Given $b \in \mathcal{A}$, by Proposition 11.5.23 we have $b = \lim ab_n a$, so

$$|\psi_j(b)| \leq |\psi_j(ab_na)| + |\psi_j(b - ab_na)| \leq |\psi_j(ab_na)| + \|b - ab_na\|.$$

It follows that $\limsup_j |\psi_j(b)| \leq \|b - ab_na\|$; and as we are free to choose $n$, we get by the Limsup Routine that $\lim_j \psi_j(b) = 0$. That is, $\psi_j \xrightarrow{\text{weak}^*} 0$, and so $0 \in \overline{S(\mathcal{A})}^{w^*}$. This is not in itself a proof that $0 \in \overline{S(\mathcal{A})}^{w^*}$, but rather a contradiction with the assumption $0 \notin \overline{S(\mathcal{A})}^{w^*}$; that is, we proved a statement of the form $P \implies \sim P$, which implies that $P$ is false. It follows that $0 \in \overline{S(\mathcal{A})}^{w^*}$.

Now that we have that $0 \in \overline{S(\mathcal{A})}^{w^*}$, given any $\psi \in \mathcal{A}^*$ with $\psi \geq 0$ and $\|\psi\| \leq 1$, let $\beta = \|\psi\|$. Then $\frac{1}{\beta}\,\psi \in S(\mathcal{A})$, and as $\overline{S(\mathcal{A})}^{w^*}$ is convex,

$$\psi = (1-\beta)\,0 + \beta\left(\frac{1}{\beta}\,\psi\right) \in \overline{S(\mathcal{A})}^{w^*}. \qquad \square$$

Some intuition of how zero can be in the closure of the state space can be obtained from looking at the case $\mathcal{A} = C_0(\mathbb{R})$. If we consider the states $\varphi_n$ given by $\varphi_n(f) = f(n)$, then $\varphi_n \to 0$ pointwise, that is in the weak*-topology. Or in $\mathcal{A} = c_0$, the states $\varphi_n(x) = x_n$ converge pointwise to zero. And to consider the other canonical example of a non-unital C*-algebra, on $\mathcal{K}(\mathcal{H})$ one can consider the states $\varphi_n(T) = \langle T\xi_n, \xi_n\rangle$ for a fixed orthonormal basis $\{\xi_n\}$. Then again $\varphi_n \to 0$ in the weak*-topology.

### 11.5.1. Exercises.

(11.5.1) Show that $\phi \in \mathcal{A}^*$ is Hermitian if and only if $\phi(a^*) = \overline{\phi(a)}$ for all $a \in \mathcal{A}$.

(11.5.2) Let $\varphi \in \mathcal{A}^*$. Put

$$\varphi^*(a) = \overline{\varphi(a^*)}.$$

Show that $\varphi^* \in \mathcal{A}^*$ and that $\varphi + \varphi^*$ is Hermitian.

(11.5.3) Show that a positive linear functional is Hermitian.

(11.5.4) Let $\varphi \in \mathcal{A}^*$ be positive, and $a, b \in \mathcal{A}$ with $b \geq 0$. Show that $|\varphi(ab)| \leq \|a\|\varphi(b)$.

(11.5.5) Let $\mathcal{A}$ be a unital C$^*$-algebra and $\mathcal{A}_0 \subset \mathcal{A}$ a $*$-subalgebra with $I_{\mathcal{A}} \in \mathcal{A}_0$. Let $\varphi : \mathcal{A}_0 \to \mathbb{C}$ be linear and positive. Show that $\varphi$ is bounded, $\|\varphi\| = \varphi(I_{\mathcal{A}})$, and $\varphi$ extends to a positive $\tilde{\varphi} \in \mathcal{A}^*$ with $\|\tilde{\varphi}\| = \|\varphi\|$.

(11.5.6) Let $\mathcal{A}$ be a C$^*$-algebra and $a \in \mathcal{A}$. Show that

$$\|a\| = \max\{\varphi(a^*a)^{1/2} : \ \varphi \in S(\mathcal{A})\}.$$

(11.5.7) We outline here a proof of Proposition 11.5.16 and Corollary 11.5.17 that goes another way. So let $\mathcal{A}$ be a C$^*$-algebra and $\varphi \in \mathcal{A}^*$ a Hermitian functional.

(i) Let $K = \{\psi \in \mathcal{A}^* : \ \ \psi \geq 0 \ \text{ and } \ \|\psi\| \leq 1\}$. Show that $K$ is weak$^*$-compact and it separates points.

(ii) Use the idea in Proposition 7.2.24 to show that $\mathcal{A}$ embeds isometrically in $C(K)$ in a way that preserves positivity.

(iii) With $\rho : \mathcal{A} \to C(K)$ the embedding, show that $\varphi \circ \rho^{-1} : \rho(\mathcal{A}) \to \mathbb{C}$ extends to a bounded Hermitian functional $\tilde{\varphi}$ on $C(K)$.

(iv) Use Theorem 5.6.15 and the Jordan decomposition of a measure to conclude that there exist $\varphi^+, \varphi^- \in S(\mathcal{A})$ and non-negative scalars $\alpha, \beta$ such that $\varphi = \alpha\varphi^+ - \beta\varphi^-$.

(11.5.8) Show that when $\mathcal{A}$ is non-unital Lemma 11.5.14 still holds if we replace condition (ii) by the existence, for every $\varepsilon > 0$, of positive $x, y \in \mathcal{A}$ with $x + y \leq I_{\mathcal{A}}$ and $\varphi(x) \geq \|\varphi\| - \varepsilon$, $\psi(y) \geq \|\psi\| - \varepsilon$.

(11.5.9) Let $\mathcal{A}$ be a C$^*$-algebra and $p \in \mathcal{A}$ a projection. Show that $p\mathcal{A}p$ is a hereditary C$^*$-subalgebra of $\mathcal{A}$.

(11.5.10) Let $\mathcal{A}$ be a C$^*$-algebra and $a \in \mathcal{A}^+$. Show that $\overline{a\mathcal{A}a}$ is a hereditary C$^*$-subalgebra of $\mathcal{A}$. Show by example that the closure is needed in general.

(11.5.11) Show that the ideal generated by the identity function in $C[0,1]$ is an example of a non-closed, non-hereditary ideal of a C$^*$-algebra. Use the example to give an explanation for why the argument after Definition 11.5.19 fails in that case.

(11.5.12) Let $\mathcal{A}$ be a non-unital C$^*$-algebra and $\tilde{\mathcal{A}}$ its unitization. Show that $\psi((a, \lambda)) = \lambda$ defines a state on $\tilde{\mathcal{A}}$.

(11.5.13) Let $\mathcal{B} \subset \mathcal{A}$ be C$^*$-algebras with $\mathcal{B}$ hereditary and $\mathcal{A}$ non-unital. Show that $\mathcal{B}$ is hereditary in $\tilde{\mathcal{A}}$.

(11.5.14) Let $a \in \mathcal{A}$ positive. Show that $a$ is strictly positive if and only if $\left\{a\left(\frac{1}{n}\, I_{\mathcal{A}} + a\right)^{-1}\right\}_n$ is an approximate unit for $\mathcal{A}$.

(11.5.15) Let $a \in \mathcal{A}$ positive with $\|a\| \leq 1$. Show that $a$ is strictly positive if and only if $\{a^{1/n}\}_n$ is an approximate unit for $\mathcal{A}$.

## 11.6. Representations

As defined and constructed, C*-algebras are abstract. But their structure (adjoints, positivity, the C*-identity) reminds us of $\mathcal{B}(\mathcal{H})$; in fact, any closed $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ is a C*-algebra. This is not by chance, as we will see.

**11.6.1. Definition.**

*A* ***representation*** *of a C*-algebra $\mathcal{A}$ is a $*$-homomorphism $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$, for some Hilbert space $\mathcal{H}$.*

*The representation $\pi$ is* ***faithful*** *if it is injective.*

*A vector $\xi \in \mathcal{H}$ is* ***cyclic*** *for $\pi$ and $\mathcal{A}$ if $\overline{\pi(\mathcal{A})\xi} = \mathcal{H}$.*

*A set $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is* ***non-degenerate*** *if $\overline{\mathcal{M}\mathcal{H}} = \mathcal{H}$. A representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ is* ***non-degenerate*** *if $\overline{\pi(\mathcal{A})\mathcal{H}} = \mathcal{H}$.*

The reason we are using "faithful" instead of injective steams from the fact that this is coherent with Definition 11.5.11; indeed, because $\pi$ is a $*$-homomorphism, $\pi(a) = 0$ if and only if $\pi(a^*a) = 0$.

Though not mentioned that way, some representations have shown up already:

- the C*-algebra $M_n(\mathbb{C})$ is systematically identified with $\mathcal{B}(\mathbb{C}^n)$. The map $\pi$ that sends a matrix to its corresponding operator is a representation. Here the representation depends on the choice of an orthonormal basis. Any nonzero $\xi \in \mathcal{H}$ is cyclic for $\pi$.
- Let $\mathcal{A} = C[0,1]$, $\mathcal{H} = L^2[0,1]$, and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ the map $\pi(f) = M_f$, the multiplication operator $M_f g = fg$. The vector $\xi = 1 \in \mathcal{H}$ is cyclic, since $C[0,1]$ is dense in $L^2[0,1]$ (Proposition 2.8.18). Multiplication operators will be discussed in Section 12.2.
- The two previous examples show representations that somehow preserve $\mathcal{A}$ (properly, those representations are faithful). Here is a different flavour of an example. Let $\mathcal{A} = C[0,1]$, $\mathcal{H} = \mathbb{C}$, and $\pi(f) = f(0)$. The $\pi$ is a representation, and any nonzero $\lambda \in \mathbb{C}$ is cyclic for $\pi$.

Given an abstract C*-algebra, it is not a priori clear that representations exist. After all, how are we going to produce a Hilbert space when all we have at hand is a C*-algebra? Some inspiration can be found from the second example above. If we start with $\mathcal{A} = C[0,1]$ and consider the inner product $\langle f, g\rangle = \int_0^1 f\bar{g}$, then $\mathcal{A}$ becomes a pre-Hilbert space and its completion is $L^2[0,1]$, where $\mathcal{A}$ acts by multiplication. We will use a similar idea for an arbitrary C*-algebra; namely, given a state $\varphi \in S(\mathcal{A})$, we can get an inner product by $\langle a, b\rangle = \varphi(b^*a)$. The implementation of this idea is called the **GNS Construction**.

***11.6.2. Theorem. (Gelfand–Naimark–Segal, GNS)***

*Let $\mathcal{A}$ be a $C^*$-algebra, and $\varphi \in S(\mathcal{A})$. Then there exists a Hilbert space $\mathcal{H}_\varphi$, $\xi_\varphi \in \mathcal{H}_\varphi$, and a representation $\pi_\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\varphi)$ such that $\varphi(a) = \langle \pi_\varphi(a)\xi_\varphi, \xi_\varphi\rangle$ for all $a \in \mathcal{A}$. Moreover $\xi_\varphi$ is cyclic for $\pi$, that is $\overline{\pi_\varphi(\mathcal{A})\xi_\varphi} = \mathcal{H}_\varphi$.*

*Proof.* Consider on $\mathcal{A}$ the sesquilinear form $[a, b] = \varphi(b^*a)$. Let $L = \{a \in \mathcal{A} : \varphi(a^*a) = 0\}$; this is a subspace of $\mathcal{A}$, due to Cauchy–Schwarz. Indeed, we have that

$$\begin{aligned} L &= \{a \in \mathcal{A} :\ \varphi(ba) = 0,\ b \in \mathcal{A}\} \\ &= \{a \in \mathcal{A} :\ [b, a] = [a, b] = 0,\ b \in \mathcal{A}\}. \end{aligned} \tag{11.18}$$

This follows from $|\varphi(ba)| \le \varphi(b^*b)^{1/2}\varphi(a^*a)^{1/2}$. On the vector space $\mathcal{A}/L$, define a sesquilinear form by $\langle a + L, b + L\rangle = [a, b]$; this is well-defined and an inner product by the estimates above (Exercise 11.6.2). Let $\mathcal{H}_\varphi$ be the completion of $\mathcal{A}/L$ under the norm induced by the inner product $\langle\cdot,\cdot\rangle$.

Define $\pi_\varphi : \mathcal{A} \to \mathcal{A}/L$ by $\pi_\varphi(a)(b + L) = ab + L$. This is easily seen to be linear and multiplicative. And

$$\|\pi_\varphi(a)(b+L)\|^2 = \varphi(b^*a^*ab) \le \varphi(b^*\|a^*a\|b) = \|a^*a\|\,\varphi(b^*b) = \|a\|^2\,\|b+L\|^2.$$

So $\pi_\varphi(a)$ is bounded with $\|\pi(a)\| \le \|a\|$, and thus it extends from $\mathcal{A}/L$ to $\mathcal{H}_\varphi$, that is, $\pi_\varphi(a) \in \mathcal{B}(\mathcal{H}_\varphi)$. Also,

$$\begin{aligned} \langle \pi_\varphi(a)(b+L), c+L\rangle &= \langle ab+L, c+L\rangle = \varphi(c^*ab) \\ &= \varphi((a^*c)^*b) = \langle b+L, \pi_\varphi(a^*)(c+L)\rangle \end{aligned}$$

for all $b, c \in \mathcal{A}$, so $\pi_\varphi(a)^* = \pi_\varphi(a^*)$. Hence $\pi_\varphi$ is a representation of $\mathcal{A}$ on $\mathcal{H}_\varphi$.

If $\mathcal{A}$ is unital, put $\xi_\varphi = I_\mathcal{A} + L$. Then $\langle \pi_\varphi(a)\xi_\varphi, \xi_\varphi\rangle = \varphi(a)$. If $\mathcal{A}$ is not unital, let $\{e_j\}$ be an approximate unit. By the Riesz Representation Theorem (4.5.4) and Banach–Alaoglu (Theorem 7.2.13), the unit ball in a Hilbert space is weakly compact. So $\{e_j + L\}$ admits a weakly convergent subnet $\{e_{j_k}\}$; let $\xi_\varphi$ be its limit. Then

$$\begin{aligned} \langle \pi_\varphi(a)\xi_\varphi, \xi_\varphi\rangle &= \lim_k \lim_h \langle \pi_\varphi(a)(e_{j_k} + L), (e_{j_h} + L)\rangle \\ &= \lim_k \lim_h \varphi(e_{j_h}^* a e_{j_k}) = \lim_k \varphi(a e_{j_k}) = \varphi(a). \end{aligned}$$

Finally we have, for $a \in \mathcal{A}$,

$$\langle \pi_\varphi(a)\xi_\varphi, b+L\rangle = \lim_k \langle \pi_\varphi(a)(e_{j_k} + L), b+L\rangle = \lim_k \varphi(b^*ae_{j_k})$$

$$= \varphi(b^*a) = \langle a+L, b+L\rangle.$$

As $b \in \mathcal{A}$ was arbitrary and $\{b+L : \ b \in \mathcal{A}\}$ is dense in $\mathcal{H}_\varphi$, we get that $\pi_\varphi(a) = a+L$. Hence $\pi_\varphi(\mathcal{A})\xi_\varphi = \mathcal{A}/L$ is dense. ☐

It is more or less standard to denote the GNS representation of $\mathcal{A}$ corresponding to a state $\varphi$ by $(\pi_\varphi, \mathcal{H}_\varphi, \xi_\varphi)$. Sometimes $\Omega_\varphi$ is used to denote the vector instead of $\xi_\varphi$. An important observation is that every cyclic representation is a GNS representation, in the following sense:

***11.6.3. Proposition.***

*Let $\mathcal{A}$ be a $C^*$-algebra, $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a cyclic representation, and $\xi \in \mathcal{H}$ a unit vector such that $\overline{\pi(\mathcal{A})\xi} = \mathcal{H}$. Then $\varphi \in S(\mathcal{A})$, where $\varphi(a) = \langle \pi(a)\xi, \xi\rangle$, and there exists a unitary $U : \mathcal{H} \to \mathcal{H}_\varphi$ such that $\pi(a) = U\pi_\varphi(a)U^*$ for all $a \in \mathcal{A}$.*

*Proof.* Define $\varphi(a) = \langle \pi(a)\xi, \xi\rangle$. Since $\|\xi\| = 1$, $\varphi \in S(\mathcal{A})$. Let $U : \pi_\varphi(\mathcal{A})\xi_\varphi \to \pi(\mathcal{A})\xi$ be given by $U\pi_\varphi(a)\xi_\varphi = \pi(a)\xi$. We will see that this map is isometric, which will also guarantee that it is well-defined. We have, for $a, b \in \mathcal{A}$,

$$\begin{aligned}\langle U\pi_\varphi(a)\xi_\varphi, U\pi_\varphi(b)\xi_\varphi\rangle &= \langle \pi(a)\xi, \pi(b)\xi\rangle = \langle \pi(b^*a)\xi, \xi\rangle = \varphi(b^*a) \\ &= \langle \pi_\varphi(b^*a)\xi_\varphi, \xi_\varphi\rangle = \langle \pi_\varphi(a)\xi_\varphi, \pi_\varphi(b)\xi_\varphi\rangle.\end{aligned}$$

In particular, $\|U\pi_\varphi(a)\xi_\varphi\| = \|\pi_\varphi(a)\xi_\varphi\|$ and $U$ is both well-defined and isometric. So $U$ is linear and bounded, and it extends to $\overline{\pi_\varphi(\mathcal{A})\xi_\varphi}$ by Proposition 6.1.9. As $U$ is isometric and its range is dense, it is surjective and hence a unitary. We have, for $a, b \in \mathcal{A}$,

$$U\pi_\varphi(a)\,\pi_\varphi(b)\xi_\varphi = U\pi_\varphi(ab)\xi_\varphi = \pi(ab)\xi = \pi(a)\pi(b)\xi = \pi(a)U\pi_\varphi(b)\xi.$$

This shows that $U\pi_\varphi(a) = \pi(a)U$ on all of $\pi_\varphi(\mathcal{A})\xi_\varphi$; the density then makes the equality hold in all of $\mathcal{H}_\varphi$. As $U$ is a unitary, we get $\pi(a) = U\pi_\varphi(a)U^*$ for all $a \in \mathcal{A}$. ☐

*11.6.4. Remark*. Recall that $S(\tilde{\mathcal{A}}) = S(\mathcal{A}) \cup \{1\}$ (see Proposition 11.5.6 and the paragraph before it); so for talking about states we may assume that $\mathcal{A}$ is unital. We will consider an equivalence relation on $S(\mathcal{A})$, given by $\psi \sim \varphi$ if there exists a unitary $u \in \mathcal{A}$ such that $\psi(a) = \varphi(u^*au)$ for all $a \in \mathcal{A}$ (we need to work on the unitization to be able to talk about unitaries). This is significant because equivalent states produce equivalent GNS representation. Indeed, looking at the GNS representations we have

$$\langle \pi_\psi(a)\xi_\psi, \xi_\psi\rangle = \psi(a) = \varphi(uau^*) = \langle \pi_\varphi(uau^*)\xi_\varphi, \xi_\varphi\rangle = \langle \pi_\varphi(a)\,\tilde{\xi}, \tilde{\xi}\rangle,$$

where $\tilde{\xi} = \pi_\varphi(u^*)\xi_\varphi$. Now Proposition 11.6.3 produces a unitary $V_u : \mathcal{H}_\psi \to \mathcal{H}_\varphi$ with

$$\pi_\psi(a) = V_u^* \pi_\varphi(a) V_u, \quad a \in \mathcal{A}. \tag{11.19}$$

*11.6.5. Remark.* For a given state $\varphi$, the space $\mathcal{H}_\varphi$ from Theorem 11.6.2 can be very small, even if $\mathcal{A}$ is big. For instance, take $\mathcal{A} = c_0$, and $\varphi(a) = a(1)$, then the space $\mathcal{H}_\varphi$ constructed above is one-dimensional. But things are different when we put all states on $\mathcal{A}$ to work together.

In Definition 5.3.8 and Remark 5.3.9 we defined the $\ell^p$ direct sum of Banach spaces. In the particular case where $p = 2$, the resulting direct sum $\bigoplus_j \mathcal{H}_j$ is a Hilbert space (Exercise 11.6.1). Though formally the elements of $\bigoplus_j \mathcal{H}_j$ are functions $g$ with $g(j) \in \mathcal{H}_j$ (and the $\ell^2$-condition $\sum_j \|g(j)\|^2 < \infty$), it is common practice to write instead $\{\xi_j\}_{j\in J}$ (where, simply, $\xi_j$ is notation for $\xi(j)$).

**11.6.6. Corollary. (Faithful Representation)**

*Let $\mathcal{A}$ be a $C^*$-algebra. Then there exists a Hilbert space $\mathcal{H}$ and a representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ such that $\mathcal{A} \simeq \pi(\mathcal{A})$.*

*Proof.* Let $\mathcal{F}$ be a family of representatives in $S(\mathcal{A})$. Let $\mathcal{H} = \bigoplus_{\varphi\in\mathcal{F}} \mathcal{H}_\varphi$. Define $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ by $\pi(a) = \bigoplus_{\varphi\in\mathcal{F}} \pi_\varphi$. That is, given $\eta \in \mathcal{H}$ and $a \in \mathcal{A}$, we have $(\pi(a)\eta)(\varphi) = \pi_\varphi(a)\eta(\varphi)$. It is straightforward to check that $\pi$ is a representation of $\mathcal{A}$. And if $\pi(a) = 0$, then $\pi_\varphi(a^*a) = 0$ for all $\varphi \in \mathcal{F}$. By (11.19), $\pi_\varphi(a^*a) = 0$ for all $\varphi \in S(\mathcal{A})$; this in turn implies that $\varphi(a^*a) = 0$ for all $\varphi \in S(\mathcal{A})$, and so $a = 0$ by Corollary 11.5.8 and the C$^*$-identity. Thus $\pi$ is injective and then it is isometric and has closed range by Proposition 11.4.9. It follows that $\pi(\mathcal{A}) \subset \mathcal{B}(\mathcal{H})$ is a C$^*$-algebra and $\pi(\mathcal{A}) \simeq \mathcal{A}$. □

The proof of Corollary 11.6.6 uses that if $\pi_\varphi(a) = 0$, then $\varphi(a) = 0$; we note that the converse of the statement is not true in general (Exercise 11.6.4). Because $\pi_\varphi$ is a representation, injectivity is the same as faithfulness for it. We have that if $\varphi$ is faithful then $\pi_\varphi$ is faithful, but the converse is not necessarily true (Exercise 11.6.6).

When $\mathcal{A}$ is separable, we may avoid the construction in Corollary 11.6.6 and instead use Proposition 11.5.12 and Theorem 11.6.2 to get a faithful representation. This is a bit more than a matter of just different methods. Consider $\mathcal{A} = M_2(\mathbb{C})$. Let us see what kind of representations we get if apply

Corollary 11.6.6 versus applying GNS to a faithful state. We know from Theorem 10.7.11, after identifying $\mathcal{A}$ with $\mathcal{K}(\mathbb{C}^2)$, that states on $\mathcal{A}$ are of the form $\varphi_S : A \longmapsto \operatorname{Tr}(AS)$ for $S \in \mathcal{A}$ (the finite dimension gives us $\mathcal{T}(\mathcal{H}) = \mathcal{B}(\mathcal{H})$). The positivity of $\varphi_S$ forces $S \geq 0$, as for any $\xi \in \mathcal{H}$ we have $\langle S\xi, \xi\rangle = \varphi(P)$, where $P$ is the rank-one projection onto $\mathbb{C}\xi$. And $1 = \|\varphi_S\| = \varphi(I_2) = \operatorname{Tr}(S)$. So the state space of $\mathcal{A}$ can be identified with the positive elements of $\mathcal{A}$ that have trace one. The unitary conjugacy class of $S$ can be identified with the eigenvalues $\lambda_1, \lambda_2$, which are non-negative and add to 1; thus $\lambda_1$ determines $\lambda_2 = 1 - \lambda_1$. In the notation of Corollary 11.6.6, we can identify $\mathcal{F}$ with $[0, 1]$, where

$$\varphi_t(A) = \operatorname{Tr}\left( \begin{bmatrix} t & 0 \\ 0 & 1-t \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right) = ta_{11} + (1-t)a_{22}.$$

When $t \in (0, 1)$ the state $\varphi_t$ is faithful. This means that the GNS space $\mathcal{H}_t$ is linearly isomorphic to $\mathcal{A}$ ($L = \{0\}$ in the notation of Theorem 11.6.2) and so $\dim \mathcal{H}_t = \dim \mathcal{A} = 4$. The inner product is

$$\langle A, B\rangle = \varphi_t(B^*A) = t(a_{11}\overline{b_{11}} + a_{21}\overline{b_{21}}) + (1-t)(a_{12}\overline{b_{12}} + a_{22}\overline{b_{22}})$$

and $\pi_t(A)B = AB$. When $t \in (0, 1)$, and we do GNS for $\varphi_t$, $\mathcal{H}_t = M_2(\mathbb{C})$ (since $L_t = \{0\}$), and $\{t^{-1/2}\, E_{11}, (1-t)^{-1/2}\, E_{12}, t^{-1/2}\, E_{21}, (1-t)^{-1/2}\, E_{22}\}$ is an orthonormal basis for $\mathcal{H}_t$. With respect to this basis

$$\pi_t(A) = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & a_{21} & 0 & a_{22} \end{bmatrix}.$$

For instance the $2, 4$ entry is (writing $E_{22}$ instead of $E_{22} + L$, etc., to simplify notation)

$$\begin{aligned} \langle \pi(A)(1-t)^{-1/2}E_{22}, (1-t)^{-1/2}E_{12}\rangle &= (1-t)^{-1}\varphi_t(E_{21}AE_{22}) \\ &= (1-t)^{-1}(1-t)a_{12} = a_{12}. \end{aligned}$$

When $t = 1$, instead, the subspace $L$ consists of those $A$ such that $(A^*A)_{11} = 0$, which are the matrices of the form

$$A = \begin{bmatrix} 0 & a_{12} \\ 0 & a_{22} \end{bmatrix}.$$

So in this case $\dim \mathcal{H}_1 = 2$ and similarly $\dim \mathcal{H}_0 = 2$ (and both $\pi_1$ and $\pi_0$ are faithful! This, since any nonzero representation on $M_2(\mathbb{C})$ is faithful by Exercise 11.6.5). In both the $t = 1$ and $t = 0$ cases we recover the canonical representation of $\mathcal{A}$ as operators on $\mathbb{C}^2$. Indeed, if we look at the case $t = 1$, an orthonormal basis of $\mathcal{H}_1$ is given by $\{E_{11}, E_{21}\}$, and with respect to this orthonormal basis the matrix of $A \in \mathcal{A}$ is $A$. For $t = 0$ the canonical orthonormal basis is $\{E_{12}, E_{22}\}$.

As opposed to the fact that each $\pi_t$ gives faithful representations of $\mathcal{A}$ on a finite-dimensional Hilbert space (with $\pi_0$ and $\pi_1$ of dimension 2), $\mathcal{H}$ from

Corollary 11.6.6 is infinite-dimensional. Basically $\pi(A)$ consists of copying $A$ infinitely many times, plus one copy of $a_{11}$ and one of $a_{22}$.

The proof of Corollary 11.6.6 uses a concept that deserves to be given a name. And that is the idea of adding representations as orthogonal direct sums. Namely,

**11.6.7. Definition.**

*Let $\mathcal{A}$ be a $C^*$-algebra and $\{\pi_j\}$ a family of representations of $\mathcal{A}$, with $\pi_j : \mathcal{A} \to \mathcal{B}(\mathcal{H}_j)$ for each $j$. The* ***direct sum*** *(or, simply,* ***sum****) is the representation $\bigoplus \pi_j : \mathcal{A} \to \bigoplus_h \mathcal{H}_j$ given by*

$$\Big(\bigoplus_j \pi_j\Big)(a) = \bigoplus_j \pi_j(a).$$

***11.6.8. Proposition.***

*Let $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a non-degenerate representation. Then $\pi$ is a direct sum of cyclic representations. When $\pi$ is degenerate, it can be written as a sum of the zero representation and a direct sum of cyclic representations.*

*Proof.* Let $\mathcal{F}$ be the collection of families $\{\xi_k\}_k$ of vectors in $\mathcal{H}$ such that $\{\pi(\mathcal{A})\xi_k\}$ are pairwise orthogonal. The collection $\mathcal{F}$ is nonempty because it contains the one-element families $\{\xi\}$ for nonzero $\xi \in \mathcal{H}$. We consider inclusion as the order for $\mathcal{F}$. Given a chain of families, the union is an upper bound. Hence by Zorn's Lemma there exists a maximal family $\{\xi_j\}$ in $\mathcal{F}$. Write $\mathcal{H}_j = \overline{\pi(\mathcal{A})\xi_j}$ and $\pi_j : \mathcal{A} \to \mathcal{B}(\mathcal{H}_j)$ the representation defined by $\pi_j(a)\pi(b)\xi_j = \pi(ab)\xi_j$ and extended by continuity. If $\xi \in \left(\bigoplus_j \mathcal{H}_j\right)^\perp$, then for any $j$ we have

$$\langle \pi(a)\xi, \pi(b)\xi_j \rangle = \langle \xi, \pi(a^*b)\xi_j \rangle = 0,$$

so $\pi(\mathcal{A})\xi \perp \pi(\mathcal{A})\xi_j$ for all $j$, and the family $\{\xi_j\} \cup \{\xi\}$ is in $\mathcal{F}$; but this contradics the maximality. Thus $\mathcal{H} = \bigoplus_j \mathcal{H}_j$, and $\pi = \bigoplus_j \pi_j$. $\square$

Another consequence of having GNS representations available is the following description of linear functionals.

***11.6.9. Proposition.***

*Let $\mathcal{A}$ be a C*-algebra and $\varphi \in \mathcal{A}^*$. Then there exists a state $\psi \in S(\mathcal{A})$ and $\xi, \eta \in \mathcal{H}_\psi$ such that*

$$\varphi(a) = \langle \pi_\psi(a)\xi, \eta \rangle, \qquad a \in \mathcal{A}.$$

*Proof.* Via Proposition 11.5.16 we can write $\varphi = \psi_1 - \psi_2 + i(\psi_3 - \psi_4)$ with $\psi_1, \psi_2, \psi_3, \psi_4$ positive linear functionals. Let $\psi' = \psi_1 + \psi_2 + \psi_3 + \psi_4$ and $\psi = \frac{1}{\|\psi'\|}\,\psi'$. Then $\psi \in S(\mathcal{A})$ and $\psi_j \leq \psi'$ for all $j = 1, \ldots, 4$. Let $\tilde{\psi}_1 : \pi_\psi(\mathcal{A})\xi_\psi \to \mathbb{C}$ be given by $\tilde{\psi}_1(\pi(a)\xi_\psi) = \psi_1(a)$. This is well-defined, for if $\pi(a)\xi_\psi = 0$ then $\psi(a^*a) = 0$ and hence

$$|\tilde{\psi}_1(\pi(a)\xi_\psi)| = |\psi_1(a)|^2 \leq \psi_1(a^*a) \leq \psi'(a^*a) = \|\psi'\|\,\psi(a^*a) = 0.$$

And it is bounded, for

$$\begin{aligned} |\tilde{\psi}_1(\pi(a)\xi_\psi)| &= |\psi_1(a)| \leq \psi_1(a^*a)^{1/2} \\ &\leq \psi'(a^*a)^{1/2} = \|\psi'\|^{1/2}\,\psi(a^*a)^{1/2} = \|\psi'\|^{1/2}\,\|\pi_\psi(a)\xi_\psi\|. \end{aligned}$$

Therefore $\tilde{\psi}_1$ extends to a bounded linear functional on $\mathcal{H}_\psi$. By the Riesz Representation Theorem (4.5.4) there exists $\eta_1 \in \mathcal{H}_\psi$ such that

$$\psi_1(a) = \langle \pi_\psi(a)\xi_\psi, \eta_1 \rangle, \qquad a \in \mathcal{A}.$$

Repeating the same argument for the other three functionals and putting $\eta = \eta_1 - \eta_2 - i(\eta_3 - \eta_4)$ we get

$$\varphi(a) = \langle \pi_\psi(a)\xi_\psi, \eta \rangle, \qquad a \in \mathcal{A}. \qquad \square$$

**11.6.1. Exercises.**

(11.6.1) Let $\{\mathcal{H}_j\}$ be a family of Hilbert spaces and let $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ the $\ell^2$ direct sum as in Definition 5.3.8. Show that $\mathcal{H}$ is a Hilbert space.

(11.6.2) Show that in the proof of Theorem 11.6.2 the form $\langle a+L, b+L \rangle = [a, b]$ is well-defined and an inner product.

(11.6.3) If $\varphi, \psi \in S(\mathcal{A})$ with $\varphi \sim \psi$, construct $V$ and show the equality (11.19).

(11.6.4) Find an example of $\mathcal{A}$, $a \in \mathcal{A}$, $\varphi \in S(\mathcal{A})$, such that $\varphi(a) = 0$ but $\pi_\varphi(a) \neq 0$.

(11.6.5) Let $\mathcal{A}$ be a simple C*-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a nonzero representation. Show that $\pi$ is faithful.

(11.6.6) Let $\varphi \in S(\mathcal{A})$, $a \in \mathcal{A}$. Show that $\pi_\varphi(a) = 0$ if and only if $\varphi(b^*a^*ab) = 0$ for all $b \in \mathcal{A}$. Conclude that $\varphi$ faithful implies $\pi_\varphi$ faithful. Show that the converse is not necessarily true.

(11.6.7) Let $\varphi \in S(\mathcal{A})$ be faithful, and $a \in \mathcal{A}$. Show that if $\pi_\varphi(a)\xi_\varphi = 0$, then $a = 0$. *We will give a name to this property of $\xi_\varphi$ in Section 12.5:* ***separating****.*

(11.6.8) Let $\mathcal{A}$ be a C*-algebra, $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation, and $\xi \in \mathcal{H}$. Let $P \in \mathcal{B}(\mathcal{H})$ be the orthogonal projection onto $\overline{\pi(\mathcal{A})\xi}$. Show that $P\pi(a) = \pi(a)P$ for all $a \in \mathcal{A}$.

(11.6.9) Let $\mathcal{A}$ be a C*-algebra and $a \in \mathcal{A}$. Show that there exists a unique $b \in \mathcal{A}$ such that $a = bb^*b$.

(11.6.10) Let $\mathcal{A}$ be a non-unital C*-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. Prove that there exists a (unique, if $\pi$ is non-degenerate) representation $\tilde{\pi} : \tilde{\mathcal{A}} \to \mathcal{B}(\mathcal{H})$ that extends $\pi$. If $\pi$ is faithful, so is $\tilde{\pi}$.

## 11.7. Matrices over a C*-algebra

Given a C*-algebra $\mathcal{A}$ and $n \in \mathbb{N}$, we consider $M_n(\mathcal{A})$. This can be made into a $*$-algebra by defining the operations the same way we do for matrices. Namely, given $a, b \in M_n(\mathcal{A})$, $a + b$ is the matrix with entries $(a + b)_{kj} = a_{kj} + b_{kj}$, and $ab$ is the matrix with entries $(ab)_{kj} = \sum_h a_{kh} b_{hj}$. Similarly, $a^*$ is the matrix with entries $(a^*)_{kj} = a^*_{jk}$. What is less obvious is how to make $M_n(\mathcal{A})$ into a C*-algebra; that is, we want to define a C*-norm on it.

By Corollary 11.6.6 we may assume that $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. This allows us to think of $M_n(\mathcal{A})$ as a subalgebra of $\mathcal{B}(\mathcal{H}^n)$. Namely, given $a \in M_n(\mathcal{A})$ and $\tilde{\xi} \in \mathcal{H}^n$, we define

$$a\tilde{\xi} = \Big(\sum_{j=1}^n a_{kj}\xi_j\Big)_{k=1}^n \in \mathcal{H}^n.$$

It's clear that this makes $a$ a bounded operator, for

$$\|a\tilde{\xi}\|^2 = \sum_{k=1}^n \Big\|\sum_{j=1}^n a_{kj}\xi_j\Big\|^2 \leq \sum_{k=1}^n \Big(\sum_{j=1}^n \|a_{kj}\|\,\|\xi_j\|\Big)^2 \leq \sum_{k=1}^n \Big(\sum_{j=1}^n \|a_{kj}\|^2\Big)\,\|\tilde{\xi}\|^2,$$

so $a$ is bounded. We define a norm on $M_n(\mathcal{A})$ as the operator norm in the above picture, that is

$$\|a\| = \sup\{\|a\tilde{\xi}\| : \ \tilde{\xi} \in \mathcal{H}^n,\ \|\tilde{\xi}\| = 1\}.$$

As the norm is an operator norm, it satisfies the C*-identity. Next we claim that $M_n(\mathcal{A})$ is complete under this norm, which then says that $M_n(\mathcal{A})$ is a C*-algebra; and, in particular, it admits a single C*-norm (as mentioned on page 757). To see that $M_n(\mathcal{A})$ is complete, suppose that $\{a^{(m)}\} \subset M_n(\mathcal{A})$ is Cauchy. Then for each $k, j \in \{1, \ldots, n\}$, if $\xi \in \mathcal{H}$ and $\tilde{\xi} = (0, \ldots, 0, \xi, 0, \ldots, 0)$, with $\xi$ in the $j^{\text{th}}$ entry,

$$\|(a^{(m)}_{kj} - a^{(r)}_{kj})\xi\| = \|(a^{(m)} - a^{(r)})\tilde{\xi}\| \leq \|(a^{(m)} - a^{(r)})\|\,\|\tilde{\xi}\| = \|(a^{(m)} - a^{(r)})\|\,\|\xi\|,$$

so $\{a_{kj}^{(m)}\}$ is Cauchy in $\mathcal{A}$. Then there exists $a_{kj} = \lim_m a_{kj}^{(m)} \in \mathcal{A}$. Now fix $\varepsilon > 0$. Let $m_0$ such that for all $m \geq m_0$ and all $k, j$, $\|a_{kj} - a_{kj}^{(m)}\| < \varepsilon^{1/2}/n$. We have, for arbitrary $\tilde{\xi} \in \mathcal{H}^n$,

$$\|(a - a^{(m)})\tilde{\xi}\|^2 = \sum_{k=1}^n \left\| \sum_{j=1}^n (a_{kj} - a_{kj}^{(m)})\xi_j \right\|^2 \leq \sum_{k=1}^n \Big( \sum_{j=1}^n \|a_{kj} - a_{kj}^{(m)}\| \, \|\xi_j\| \Big)^2$$

$$\leq \sum_{k=1}^n \Big( \sum_{j=1}^n \|a_{kj} - a_{kj}^{(m)}\|^2 \Big) \sum_{j=1}^n \|\xi_j\|^2 \leq \varepsilon \, \|\tilde{\xi}\|.$$

That is, $\|a - a^{(m)}\| \to 0$. It follows from these estimates that we always have

$$\max\{\|a_{kj}\| : \ k, j\} \leq \|a\| \leq n^2 \max\{\|a_{kj}\| : \ k, j\}. \tag{11.20}$$

So convergence in $M_n(\mathcal{A})$ is precisely entry-wise convergence.

We mention that it is possible to have a completely different approach to defining $M_n(\mathcal{A})$, which is to define it as the tensor product $M_n(\mathbb{C}) \otimes \mathcal{A}$ (see Section 13.1). The net result is the same, though, as $M_n(\mathcal{A})$ becomes a C*-algebra and there is a single possible C*-norm on it. A big advantage of this latter approach is notational. Namely, we can use $E_{kj} \otimes a$ to denote the matrix with $a$ in the $k, j$ entry and zeroes elsewhere. This allows us to write $\tilde{a} \in M_n(\mathcal{A})$ as $\tilde{a} = \sum_{k,j} E_{kj} \otimes a_{kj}$, and this notation behaves nicely with algebraic operations, for instance

$$(E_{kj} \otimes a)(E_{st} \otimes b) = E_{kj}E_{st} \otimes ab) = \delta_{s,j} \, E_{kt} \otimes ab.$$

Let us first collect some basic results:

**11.7.1. Proposition.**

*Let $\mathcal{A}$ be a C*-algebra and $n \in \mathbb{N}$.*

(i) *If $a_1 \ldots, a_n \in \mathcal{A}$, then* $\left\| \sum_{k=1}^n E_{kk} \otimes a_k \right\| = \max\{\|a_k\| : \ k = 1, \ldots, n\}$.

(ii) *Let $a \in \mathcal{A}$. Then* $\left\| \begin{bmatrix} 0 & a \\ a^* & 0 \end{bmatrix} \right\| = \|a\|$.

(iii) *Let $a, b \in \mathcal{A}$. Then* $\begin{bmatrix} I_{\mathcal{A}} & a \\ a^* & b \end{bmatrix} \geq 0$ *if and only if* $a^*a \leq b$.

(iv) *If* $a = \sum_{k,j=1}^n E_{kj} \otimes a_{kj}$, *then* $\|a_{kj}\| \leq \|a\|$ *for all* $k, j$.

*Proof.* Exercise 11.7.1. $\square$

Now we move to slightly less trivial estimates.

**11.7.2. Lemma.**

*If $a, b \in \mathcal{A}$, then $A = \begin{bmatrix} I & a \\ a^* & b \end{bmatrix} \geq 0$ if and only if $a^*a \leq b$.*

*Proof.* If $A \geq 0$, then for any $\xi \in H$,

$$\langle (b - a^*a)\xi, \xi \rangle = \left\langle \begin{bmatrix} I & a \\ a^* & b \end{bmatrix} \begin{bmatrix} -a\xi \\ \xi \end{bmatrix}, \begin{bmatrix} -a\xi \\ \xi \end{bmatrix} \right\rangle \geq 0,$$

so $a^*a \leq b$. Conversely, if $a^*a \leq b$, then for any $\xi, \eta \in H$,

$$\begin{aligned} \left\langle \begin{bmatrix} I & a \\ a^* & b \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \begin{bmatrix} \xi \\ \eta \end{bmatrix} \right\rangle &= \|\xi\|^2 + \langle b\eta, \eta \rangle + 2\mathrm{Re}\,\langle a\eta, \xi \rangle \\ &\geq \|\xi\|^2 + \langle b\eta, \eta \rangle - 2\langle a^*a\eta, \eta \rangle^{1/2} \|\xi\| \\ &\geq \|\xi\|^2 + \langle b\eta, \eta \rangle - 2\langle b\eta, \eta \rangle^{1/2} \|\xi\| \\ &= (\|\xi\| - \langle b\eta, \eta \rangle^{1/2})^2 \geq 0. \end{aligned}$$

□

Here is a more matricial proof of Lemma 11.7.2:

$$\begin{aligned} \begin{bmatrix} I & a \\ a^* & b \end{bmatrix} \geq 0 &\iff \begin{bmatrix} I & 0 \\ -a^* & I \end{bmatrix} \begin{bmatrix} I & a \\ a^* & b \end{bmatrix} \begin{bmatrix} I & -a \\ 0 & I \end{bmatrix} \geq 0 \\ &\iff \begin{bmatrix} I & 0 \\ 0 & b - a^*a \end{bmatrix} \geq 0 \iff b - a^*a \geq 0 \iff a^*a \leq b. \end{aligned}$$

**11.7.3. Lemma.**

*Let $a, b \in \mathcal{A}$. If $\begin{bmatrix} b & a \\ a^* & b \end{bmatrix} \geq 0$, then $a^*a \leq \|b\|\, b$. In particular, $\|a\| \leq \|b\|$.*

*Proof.* Note first that $b \geq 0$, since for every $\xi \in H$,

$$0 \leq \left\langle \begin{bmatrix} b & a \\ a^* & b \end{bmatrix} \begin{bmatrix} \xi \\ 0 \end{bmatrix}, \begin{bmatrix} \xi \\ 0 \end{bmatrix} \right\rangle = \langle b\xi, \xi \rangle.$$

Then we can get this estimate that we will need momentarily:

$$\langle ba^*a\eta, \eta \rangle = \|b^{1/2}a\eta\|^2 \leq \|b^{1/2}\|^2\, \|a\eta\|^2 = \|b\|\, \langle a^*a\eta, \eta \rangle.$$

We have, for all $\eta \in H$, $t \in \mathbb{R}$

$$0 \leq \left\langle \begin{bmatrix} b & a \\ a^* & b \end{bmatrix} \begin{bmatrix} -a\eta \\ t\eta \end{bmatrix}, \begin{bmatrix} -a\eta \\ t\eta \end{bmatrix} \right\rangle = \langle ba\eta, a\eta \rangle + t^2\, \langle b\eta, \eta \rangle - 2t\langle a^*a\eta, \eta \rangle$$

As this is positive for every $t \in \mathbb{R}$, we get that

$$\langle a^* a\eta, \eta\rangle^2 \le \langle ba\eta, a\eta\rangle \langle b\eta, \eta\rangle \le \|b\| \langle a^* a\eta, \eta\rangle \langle b\eta, \eta\rangle.$$

Thus, $\langle a^* a\eta, \eta\rangle \le \|b\| \langle b\eta, \eta\rangle$ for all $\eta$ (the case where $a\eta = 0$ needs to be treated separately, but gives no hassle) , and so $a^* a \le \|b\| \, b$. Finally, $\|a\|^2 = \|a^* a\| \le \| \, \|b\| \, b\| = \|b\|^2$. ☐

Let us prove a strengthening of Lemma 11.7.2.

## *11.7.4. Proposition.*

*Let $a, b, c \in \mathcal{A}$. The following statements are equivalent:*

*(1)* $\begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \in M_2(\mathcal{A})^+$;

*(2) for each $\varepsilon > 0$ there exists $x_\varepsilon \in \mathcal{A}$, contractive, such that $b = (a + \varepsilon I)^{1/2} x_\varepsilon (c + \varepsilon I)^{1/2}$.*

*Proof.* (1) $\implies$ (2). We assume without loss of generality that $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$. For any $\xi \in H$,

$$\langle a\xi, \xi\rangle = \left\langle \begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \begin{bmatrix} \xi \\ 0 \end{bmatrix}, \begin{bmatrix} \xi \\ 0 \end{bmatrix} \right\rangle \ge 0.$$

So $a \ge 0$. Similarly, $c \ge 0$. Let $x_\varepsilon = (a + \varepsilon I)^{-1/2} b (c + \varepsilon I)^{-1/2}$, $a_\varepsilon = (a + \varepsilon I)^{-1/2}$, $c_\varepsilon = (c + \varepsilon I)^{-1/2}$. We have

$$\begin{bmatrix} I & x_\varepsilon \\ x_\varepsilon^* & I \end{bmatrix} = \begin{bmatrix} a_\varepsilon & 0 \\ 0 & c_\varepsilon \end{bmatrix} \begin{bmatrix} a + \varepsilon I & b \\ b^* & c + \varepsilon I \end{bmatrix} \begin{bmatrix} a_\varepsilon & 0 \\ 0 & c_\varepsilon \end{bmatrix} \ge 0.$$

By Lemma 11.7.3, $\|x_\varepsilon\| \le 1$.

(2) $\implies$ (1). Fix $\varepsilon > 0$. We can write

$$\begin{aligned} \begin{bmatrix} a + \varepsilon I & b \\ b^* & c + \varepsilon I \end{bmatrix} &= \begin{bmatrix} a + \varepsilon I & (a + \varepsilon I)^{1/2} x_\varepsilon (c + \varepsilon I)^{1/2} \\ (c + \varepsilon I)^{1/2} x_\varepsilon^* (a + \varepsilon I)^{1/2} & c + \varepsilon I \end{bmatrix} \\ &= Y \begin{bmatrix} I & x_\varepsilon \\ x_\varepsilon^* & I \end{bmatrix} Y^*, \end{aligned}$$

where

$$Y = \begin{bmatrix} (a + \varepsilon I)^{1/2} & 0 \\ 0 & (c + \varepsilon I)^{1/2} \end{bmatrix}.$$

Since $\|x_\varepsilon\| \le 1$, we have $\begin{bmatrix} I & x_\varepsilon \\ x_\varepsilon^* & I \end{bmatrix} \ge 0$ by Lemma 11.7.2, and thus

$$\begin{bmatrix} a + \varepsilon I & b \\ b^* & c + \varepsilon I \end{bmatrix} \ge 0$$

for all $\varepsilon > 0$, which gives (1). ☐

If $\mathcal{A}$ is finite-dimensional, we can improve a little:

**11.7.5. Proposition.**

*Let $a, b, c \in \mathcal{A}$, with $\mathcal{A}$ finite-dimensional. The following statements are equivalent:*

*(1)* $\begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \in M_2(\mathcal{A})^+$;

*(2) there exists $x \in \mathcal{A}$, contractive, such that $b = a^{1/2}xc^{1/2}$.*

*Proof.* (1) $\implies$ (2). We repeat the proof as in Proposition 11.7.4, for each $\varepsilon = 1/n$, $n \in \mathbb{N}$. This way we produce a sequence of contractions $\{x_n\}$ with

$$b = \left(a + \frac{1}{n}I\right)^{1/2} x_n \left(c + \frac{1}{n}I\right)^{1/2}.$$

As $\mathcal{A}$ is finite-dimensional, the unit ball is compact; thus the sequence $\{x_n\}$ has a cluster point $x$. If follows that $b = a^{1/2}xc^{1/2}$.

(2) $\implies$ (1). Here we can repeat the exact argument in the proof of Proposition 11.7.4, with $\varepsilon = 0$. $\square$

Here is a result about lifts of $*$-homomorphisms that uses matrices over a C$^*$-algebra in the proof.

**11.7.6. Proposition.**

*Let $\mathcal{A}, \mathcal{B}$ be C$^*$-algebras and $\pi : \mathcal{A} \to \mathcal{B}$ a surjective $*$-homomorphism. Then, for any $b \in \mathcal{B}$, there exists $a \in \mathcal{A}$ such that $\pi(a) = b$ and $\|a\| = \|b\|$.*

*Proof.* We may assume without loss of generality that $\|b\| = 1$. Suppose initially that $b = b^*$. As $\pi$ is surjective there exists $a \in \mathcal{A}$ with $\pi(a) = b$. By replacing $a$ with $\operatorname{Re} a$, we may assume that $a = a^*$. Define $g : \mathbb{R} \to \mathbb{R}$ to be the continuous function $g(t) = (t \wedge 1) \vee (-1)$. That is, $g(t) = t$ on $[-1, 1]$ and $g(t) = 1$ for $t \geq 1$, $g(t) = -1$ for $t < -1$. As $g$ is continuous, by Exercise 11.2.7 we have $\pi(g(a)) = g(\pi(a)) = g(b) = b$, since $g$ is the identity on $\sigma(b) \subset [-1, 1]$. We also have

$$\|b\| = \|\pi(g(a))\| \leq \|g(a)\| \leq \|g\|_\infty = \|b\|,$$

so $\|g(a)\| = \|b\|$.

When $b$ is not necessarily selfadjoint, we do the following. Let $\pi_2 : M_2(\mathcal{A}) \to M_2(\mathcal{B})$ be the 2-amplification of $\pi$, that is

$$\pi_2\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} \pi(a_{11}) & \pi(a_{12}) \\ \pi(a_{21}) & \pi(a_{22}) \end{bmatrix}.$$

It is straightforward to check that $\pi_2$ is a surjective $*$-homomorphism. We can apply the first part of the proof to the selfadjoint element

$$\tilde{b} = \begin{bmatrix} 0 & b \\ b^* & 0 \end{bmatrix}.$$

We have, using Proposition 11.7.1,

$$\|\tilde{b}\|^2 = \|\tilde{b}^*\tilde{b}\| = \left\| \begin{bmatrix} b^*b & 0 \\ 0 & bb^* \end{bmatrix} \right\| = \|b^*b\| = \|b\|^2.$$

By the first part of the proof there exists $\tilde{a} \in M_2(\mathcal{A})$ with $\pi_2(\tilde{a}) = \tilde{b}$ and $\|\tilde{a}\| = \|\tilde{b}\| = \|b\|$. If $a$ denotes the $1,2$ entry of $\tilde{a}$, we have $\pi(a) = b$, and (using Proposition 11.7.1 again)

$$\|b\| = \|\pi(a)\| \leq \|a\| \leq \|\tilde{a}\| = \|\tilde{b}\| = \|b\|.$$

Hence $\|a\| = \|b\|$. $\square$

As mentioned after Theorem 11.4.8, when $\mathcal{J} \subset \mathcal{A}$ is an ideal and $a \in \mathcal{A}$, there exists $c \in \mathcal{J}$ such that $\|a + c\| = \|a + \mathcal{J}\|$. Indeed, Proposition 11.7.6 tells us that given $a + \mathcal{J} \in \mathcal{A}/\mathcal{J}$, there exists $b \in \mathcal{A}$ such that $b + \mathcal{J} = a + \mathcal{J}$ and $\|b\| = \|a + \mathcal{J}\|$. If we take $c = b - a \in \mathcal{J}$, then $\|a + c\| = \|a + \mathcal{J}\|$.

**11.7.1. Exercises.**

(11.7.1) Prove Proposition 11.7.1 (block matrices were already considered in Section 10.4).

(11.7.2) Given a compact Hausdorff space $T$, show that the C*-algebras $\mathcal{A} = M_n(C(T))$ and $\mathcal{B} = C(T, M_n(\mathbb{C}))$ are canonically isomorphic, where the norm in $\mathcal{B}$ is given by

$$\|y\|_{\mathcal{B}} = \sup\{\|y(t)\| : \ t \in T\}.$$

(11.7.3) Let $\mathcal{A}$ be a C*-algebra and $\tilde{\mathcal{J}} \subset M_n(\mathcal{A})$. Show that $\tilde{\mathcal{J}}$ is an ideal if and only if $\tilde{\mathcal{J}} = M_n(\mathcal{J})$ for an ideal $\mathcal{J} \subset \mathcal{A}$.

## 11.8. Finite-Dimensional C*-algebras

While the classification of (separable) infinite-dimensional C*-algebras is an active area of research and a complete classification hopeless, the situation is very different in the finite-dimensional case. We will fully characterize finite-dimensional C*-algebras in Theorem 11.8.10. Some of the steps are a bit more convoluted than they could be, since we don't have von Neumann algebras and their techniques at our disposal. The ideas in the proofs should be fairly simple to anyone with some experience with von Neumann algebras. While the path we follow in this section is to work with abstract finite-dimensional C*-algebras, an alternative approach would be to first note that a finite-dimensional C*-algebra can be faithfully represented on a finite-dimensional Hilbert space and proceed from there, where the process consists now on characterizing the $*$-subalgebras of $M_n(\mathbb{C})$. The arguments would not be too different, though, with the exception that one could use the Spectral Theorem (Theorem 10.6.12) instead of proving 11.8.3.

We have already talked about projections as concrete operators in Section 10.5. Now we consider an abstract version.

**11.8.1. Definition.**

*An element $p \in \mathcal{A}$ is a* ***projection*** *if $p^*p = p$. This is usually phrased in the two equalities $p^2 = p$ and $p^* = p$.*

*We say that two projections $p, q$ are* ***pairwise orthogonal*** *if $pq = 0$. A projection $p \in \mathcal{A}$ is* ***minimal*** *if $q \leq p$, with $q$ a projection, implies that $q = 0$ or $q = p$.*

When $\mathcal{A}$ is unital, if $p$ is a projection so is $I_{\mathcal{A}} - p$, and $p$ and $I_{\mathcal{A}} - p$ are pairwise orthogonal. It is worth noting that a C*-algebra may have very few projections. For instance $C[0,1]$ only has 0 and 1 as projections and $C_0(\mathbb{R})$ only has 0. Meanwhile, $\mathcal{B}(\mathcal{H})$ has lots and lots of projections, and so do $\ell^\infty(\mathbb{N})$ and $L^\infty[0,1]$.

It will follow soon that a projection $p \in \mathcal{A}$ is minimal, in the finite-dimensional case, if and only if $p\mathcal{A}p = \mathbb{C}p$. This is not true for arbitrary C*-algebras, though. For example if $\mathcal{A} = M_2(C[0,1])$, then $p = E_{11}$ is a minimal projection but $p\mathcal{A}p \simeq C[0,1]$.

***11.8.2. Lemma.***

*Every finite-dimensional C*-algebra is unital.*

*Proof.* By Theorem 5.2.9, the unit ball of $\mathcal{A}$ is compact. By Theorem 11.4.4, the unit ball of $\mathcal{A}$ contains an approximate unit $\{e_j\}$, and because $\mathcal{A}$ is separable (Exercise 5.2.1) we may assume that the approximate unit is countable by Exercise 11.4.3 (this is not really needed, as we could have just worked with nets instead of sequences). By compactness, $\{e_j\}$ admits a convergent subsequence. The limit of such a subsequence is a unit for $\mathcal{A}$. $\square$

In a finite-dimensional C*-algebra we have a version of the spectral theorem (compare with Theorem 10.6.12):

***11.8.3. Lemma. (Finite-Dimensional Abstract Spectral Theorem)***

*Let $\mathcal{A}$ be finite-dimensional and $a \in \mathcal{A}$ normal. Then there exist scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and pairwise orthogonal projections $p_1, \ldots, p_n \in \mathcal{A}$ such that*

$$a = \sum_{j=1}^{n} \lambda_j p_j. \tag{11.21}$$

*Also, $\sigma(a) = \{\lambda_1, \ldots, \lambda_n\}$, and all elements of the spectrum are eigenvalues, in the sense that for each $j$ there exists nonzero $b \in \mathcal{A}$ such that $ab = \lambda_j b$.*

*Proof.* Suppose that $\{\lambda_j\}_{j\in\mathbb{N}} \subset \sigma(a)$ are distinct. Fix $m \in \mathbb{N}$ and consider $\lambda_1, \ldots, \lambda_m$. For each $k = 1, \ldots, m$, construct $f_k : \mathbb{C} \to \mathbb{C}$ continuous such that $f_k(\lambda_j) = 1$ if $j = 1, \ldots, k$, and $f_k(\lambda_j) = 0$ if $j > k$; while this can be done painfully by hand on the complex plane, we can use Urysohn's Lemma (Theorem 2.6.5) instead. By Functional Calculus (Definition 11.2.7), we get $f_1(a), f_2(a), \ldots, f_m(a) \in \mathcal{A}$. We will show that $f_1(a), \ldots, f_m(a)$ are linearly independent. Indeed, assume that $\sum_{j=1}^{m} \alpha_j f_j(a) = 0$. By Corollary 11.2.8, we have that $\sum_{j=1}^{m} \alpha_j f_j(\lambda_k) = 0$ for all $k = 1, \ldots, m$ (there is a subtlety here, in that we start with a linear combination of elements in $\mathcal{A}$ and then we see it as a linear combination of functions evaluated at $a$; this is no issue because the functional calculus is a homomorphism of algebras). Evaluating at $\lambda_m$, the equality is now $\alpha_m = 0$; then evaluating at $\lambda_{m-1}$ we get $\alpha_{m-1} = 0$, and so on. So $\alpha_1 = \cdots = \alpha_m = 0$ and $f_1(a), \ldots, f_m(a)$ are linearly independent. As $\mathcal{A}$ is finite-dimensional, we can have at most $m = \dim \mathcal{A}$ distinct points in $\sigma(a)$, so $\sigma(a) = \{\lambda_1, \ldots, \lambda_n\}$ with $n \leq \dim \mathcal{A}$.

Knowing that all points in the spectrum are isolated, construct continuous functions $g_j \in C(\sigma(a))$ such that $g_j(\lambda_j) = 1$ and $g_j(\lambda_k) = 0$ if $k \neq j$ (this time we can use polynomials, as there are no accumulation points on $\sigma(a)$). We have $g_j^2 = g_j$ for all $j$, and $\sum_j g_j = 1$ on $\sigma(a)$. Also, for any $t \in \sigma(a)$, $t = \sum_{j=1}^{n} \lambda_j\, g_j(t)$. Using Functional Calculus we have that $a = \sum_{j=1}^{n} \lambda_j\, g_j(a)$, and each $p_j = g_j(a)$ is a nonzero projection (note that $g_j(a)^* = \overline{g_j}(a) = g_j(a)$ since $g_j$ is real-valued and the continuous functional calculus is a $*$-homomorphism). From $g_j g_k = 0$ when $j \neq k$, we get $p_j p_k = g_j(a) g_k(a) = 0$.

Finally, given $j$ let $b = g_j(a)$. Then $a g_j(a) = \lambda_j p_j = \lambda_j g_j(a)$. $\square$

Ideals in C*-algebras may take different forms; compare for instance Proposition 7.4.13 with the ideal $\mathcal{K}(\mathcal{H})$ in $\mathcal{B}(\mathcal{H})$. In the finite-dimensional case, though, all ideals look the same:

***11.8.4. Lemma.***

*Let $\mathcal{A}$ be finite-dimensional and $\mathcal{J} \subset \mathcal{A}$ an ideal. Then there exists a central projection $z \in \mathcal{A}$ such that $\mathcal{J} = z\mathcal{A}$.*

*Proof.* As $\mathcal{J}$ is a finite-dimensional C*-algebra (Proposition 11.4.5), it has a unit $z \in \mathcal{J}^+$ by Lemma 11.8.2. Then, as $\mathcal{J}$ is an ideal and $z \in \mathcal{J}$, $z\mathcal{A} \subset \mathcal{J} = z\mathcal{J} \subset z\mathcal{A}$, so $\mathcal{J} = z\mathcal{A}$. For any $a \in \mathcal{A}$ we have $za \in \mathcal{J}$, so $za = zaz$; similarly, $az = zaz$; thus, $za = az$ and $z$ is central. $\square$

The proof of Lemma 11.8.4 shows that if an ideal $\mathcal{J} \subset \mathcal{A}$ is unital, even when $\mathcal{A}$ is infinite-dimensional, its unit $z$ is central and $\mathcal{J} = z\mathcal{A}$. This in particular shows that if $\mathcal{A}$ is infinite-dimensional and it has trivial centre, any nontrivial ideal is non-unital. An example of this situation is $\mathcal{K}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$.

This next lemma is a bit meaningless in itself; but its role is crucial in the proof of Theorem 11.8.10, where we use it to show the existence of certain partial isometries.

***11.8.5. Lemma.***

*Let $\mathcal{A}$ be finite-dimensional, such that its centre is $\mathbb{C}\, I_{\mathcal{A}}$. Let $p, q \in \mathcal{A}$ be nonzero projections. Then there exists $c \in \mathcal{A}$ with $pcq \neq 0$.*

*Proof.* Suppose that $pcq = 0$ for all $c \in \mathcal{A}$. Then $p\mathcal{A}q = \{0\}$, and so $p\mathcal{A}q\mathcal{A} = \{0\}$, where $\mathcal{A}q\mathcal{A}$ is the ideal generated by $q$. By Lemma 11.8.4 there exists a nonzero central projection $z$ with $\mathcal{A}q\mathcal{A} = z\mathcal{A}$; but as the centre of $\mathcal{A}$ is trivial, $z = I_{\mathcal{A}}$. So $\mathcal{A}q\mathcal{A} = \mathcal{A}$; this gives us $p\mathcal{A} = \{0\}$ and so $0 = p^2 = p$, a contradiction. $\square$

Now a bit of functional calculus. The following definition only works in finite dimension. In infinite dimension there are problems both with the fact that $\mathcal{A}$ may fail to contain projections, and with the fact that ranges of operators may not be closed.

**11.8.6. Definition.**

*Let $\mathcal{A}$ be a finite-dimensional C*-algebra and $a \in \mathcal{A}$. If $a$ is normal, we write $a = \sum_{j=1}^{n} \lambda_j p_j$ as in Lemma 11.8.3 (this presentation is unique if we require $\lambda_1, \dots, \lambda_n$ to be distinct and nonzero, by Exercise 11.8.5). The*

***range projection*** *of $a$ is*

$$R(a) = \sum_{j=1}^{n} p_j.$$

*For arbitrary $a \in \mathcal{A}$, we define $R(a) = R(aa^*)$.*

Here is an algebraic way to obtain $R(a)$ without knowing the spectral decomposition.

**11.8.7. Lemma.**

*Let $\mathcal{A}$ be finite-dimensional and $a \in \mathcal{A}^+$. Then $R(a) = \lim_n a\Big(a + \frac{1}{n} I_{\mathcal{A}}\Big)^{-1}$.*

*Proof.* By Lemma 11.8.3, we have that $a = \sum_{j=1}^{m} \lambda_j p_j$, with $p_1, \ldots, p_m$ pairwise orthogonal and $\lambda_1 \geq \cdots \geq \lambda_m > 0$, as nothing prevents us from omitting the zero term if present in (11.21). Then

$$a\Big(a + \frac{1}{n} I_{\mathcal{A}}\Big)^{-1} = \sum_{j=1}^{m} \frac{\lambda_j}{\lambda_j + \frac{1}{n}} p_j,$$

and

$$\begin{aligned} \left\| R(a) - a\Big(a + \frac{1}{n} I_{\mathcal{A}}\Big)^{-1} \right\| &= \left\| \sum_{j=1}^{m} \Big(1 - \frac{\lambda_j}{\lambda_j + \frac{1}{n}}\Big) p_j \right. = \frac{1}{n} \left\| \sum_{j=1}^{m} \Big(\frac{1}{\lambda_j + \frac{1}{n}}\Big) p_j \right\| \\ &= \frac{1}{n} \max\left\{ \frac{1}{|\lambda_j + \frac{1}{n}|} \,:\, j = 1, \ldots, n \right\}, \end{aligned}$$

where the last equality is due to the norm of a normal operator being equal to its spectral radius (and the spectrum is determined by Exercise 11.8.3). As there are finitely many $\lambda_j$, for $n$ big enough we have that

$$\left\| R(a) - a\Big(a + \frac{1}{n} I_{\mathcal{A}}\Big)^{-1} \right\| \leq \frac{2}{n \min\{|\lambda_j| : j\}} \xrightarrow[n\to\infty]{} 0. \qquad \square$$

We need a few more properties of projections and range projections.

**11.8.8. Lemma.**

*Let $\mathcal{A}$ be a finite-dimensional $C^*$-algebra and $a \in \mathcal{A}$. Then* *(i)* $R(|a|) = R(|a|^2) = R(a^*a) = R(a^*)$*;*

*(ii)* $aR(a^*) = a$*;*

*(iii)* $R(a^*)$ *is the smallest projection in* $\mathcal{A}$ *such that* $aR(a^*) = a$*;*

*(iv) if* $p$ *is a projection, then* $R(pa) \leq p$*.*

*Proof.*

(i) If, as in Lemma 11.8.3, $|a| = \sum_{k=1}^{n} \mu_k p_k$, then $|a|^2 = \sum_{k=1}^{n} \mu_k^2 p_k$. So
$$R(|a|) = \sum_{k=1}^{n} p_k = R(|a|^2).$$
The other two equalities occur by definition.

(ii) Again write, as in Lemma 11.8.3, $a^*a = \sum_{k=1}^{n} \lambda_k p_k$. Let
$$p = R(a^*a) = \sum_{k=1}^{n} p_k.$$
Then $a^*a(I_{\mathcal{A}} - p) = 0$. From this,
$$0 = (I_{\mathcal{A}} - p)a^*a(I_{\mathcal{A}} - p) = [a(I_{\mathcal{A}} - p)]^*a(I_{\mathcal{A}} - p),$$
and so $a(I_{\mathcal{A}} - p) = 0$. That is,
$$a = ap = aR(a^*a) = aR(a^*).$$

(iii) Suppose that $q \in \mathcal{A}$ is a projection and that $aq = a$. Then $a^*aq = a^*a$. Let $g \in \mathbb{C}[x]$ with $g|_{\sigma(a^*a)\setminus\{0\}} = 1$. Then $R(a^*) = g(a^*a)$. From $a^*aq = a^*a$ we get
$$R(a^*)q = R(a^*a)q = g(a^*a)q = g(a^*a) = R(a^*).$$
Thus $R(a^*) \leq q$ by Exercise 11.3.8.

(iv) By Lemma 11.8.3 we can write $paa^*p = \sum_j \mu_j q_j$ for distinct scalars $\mu_j > 0$ are pairwise orthogonal projections $q_j$. Then
$$(I_{\mathcal{A}} - p)q_j(I_{\mathcal{A}} - p) = \mu_j^{-1}(I_{\mathcal{A}} - p)q_j(I_{\mathcal{A}} - p) \leq (I_{\mathcal{A}} - p)paa^*p(I_{\mathcal{A}} - p) = 0.$$
Thus $0 = (I_{\mathcal{A}} - p)q_j(I_{\mathcal{A}} - p) = (I_{\mathcal{A}} - p)q_j^*q_j(I_{\mathcal{A}} - p)$ and then $q_j(I_{\mathcal{A}} - p) = 0$; that is, $q_j = pq_j = q_jp$. Applying Exercise 11.8.2 in the algebra $p\mathcal{A}p$ we get that
$$R(pa) = R(paa^*p) = \sum_j a_j \leq p. \qquad \square$$

We will need the finite-dimensional version of the Polar Decomposition; the key difference with Proposition 10.4.11 is that with $\mathcal{A}$ being finite-dimensional we can get the partial isometry to stay in $\mathcal{A}$; and that we can make everything explicit. This is not the case in general; for example $C[0,1]$ has no nontrivial partial isometries.

***11.8.9. Lemma.***

*Let $\mathcal{A}$ be finite-dimensional, $a \in \mathcal{A}$. Then there exists a partial isometry $v \in \mathcal{A}$ with $a = v|a|$ and*

$$v^*v = R(a^*), \qquad vv^* = R(a). \tag{11.22}$$

*Proof.* Using Lemma 11.8.3 we can write $|a| = \sum_{k=1}^n \lambda_k p_k$, with $\lambda_1 > \cdots > \lambda_n > 0$. Let

$$v = a\sum_{k=1}^n \lambda_k^{-1}\, p_k.$$

We compute, using Lemma 11.8.8,

$$v|a| = a\sum_{k,j=1}^n \lambda_k^{-1}\lambda_j\, p_kp_j = a\sum_{k=1}^n p_k = aR(|a|) = a.$$

We have, since $\sigma(|a|)$ is finite, that $f(t) = t^{-1}\,1_{(0,\infty)}$ is continuous on $\sigma(|a|)$. Then, using Exercise 11.2.3,

$$\begin{aligned} vv^* &= a\sum_{k,j=1}^n \lambda_k^{-1}\lambda_j^{-1}\, p_kp_j\, a^* = a\sum_{k=1}^n \lambda_k^{-2}\, p_k\, a^* = a\, f(|a|^2)\, a^* \\ &= f(|a^*|^2)aa^* = f(|a^*|^2)|a^*|^2 = R(|a^*|^2) = R(aa^*) = R(a). \end{aligned}$$

Similarly,

$$\begin{aligned} v^*v &= \sum_{k,j=1}^n \lambda_k^{-1}\lambda_j^{-1} p_k a^* a p_j = \sum_{k,j=1}^n \lambda_k^{-1}\lambda_j^{-1}\sum_{\ell=1}^n \lambda_\ell^2 p_k p_\ell p_j \\ &= \sum_{k=1}^n p_k = R(|a|) = R(a^*). \end{aligned}$$

□

**11.8.10. Theorem. (General Form of a Finite-Dimensional C*-Algebra)**

*Let $\mathcal{A}$ be a finite-dimensional C*-algebra. Then there exist $n_1, \ldots, n_k \in \mathbb{N}$ such that $\mathcal{A} \simeq \bigoplus_{j=1}^k M_{n_j}(\mathbb{C})$.*

*Proof.* Consider first the case where $\mathcal{A}$ is abelian. By Lemma 11.8.2 we have $I_\mathcal{A} \in \mathcal{A}$. And there exist pairwise orthogonal minimal projections $p_1, \ldots, p_n$ with $I_\mathcal{A} = p_1 + \cdots + p_n$. Indeed, if there was no minimal projection, we would get an ever decreasing sequence of projections $I_\mathcal{A} \geq p_1 \geq p_2 \geq \cdots$ and then the infinite set of pairwise orthogonal projections $\{p_k - p_{k+1}\}$ is linearly independent, a contradiction since $\mathcal{A}$ is finite-dimensional. Once we have $p_1$ minimal, let $p_2 \leq I_\mathcal{A} - p_2$ be minimal; then $p_1 p_2 = 0$ and $p_1 + p_2$ is a projection. Now let $p_3 \leq I_\mathcal{A} - p_1 - p_2$ be minimal. This is a finite process, again since $\mathcal{A}$ is finite-dimensional.

Now, given any $a \in \mathcal{A}^{\rm sa}$, consider $p_1 a$; as this is selfadjoint (because $\mathcal{A}$ is abelian), we have $p_1 a = \sum_{j=1}^m \lambda_j q_j$ by Lemma 11.8.3. But then $p_1 q_1$ is a projection (again because $\mathcal{A}$ is abelian) and $p_1 q_1 = p_1 q_1 p_1 \leq p_1 I_\mathcal{A} p_1 = p_1$; as $p_1$ is minimal, $p_1 q_1 = p_1$. Similarly, $p_1 q_2 = p_1$; but then $p_1 = p_1 q_1 p_1 q_2 = p_1 q_1 q_2 = 0$, a contradiction. This shows that $p_1 a = \lambda_1 p_1$. That is, since self-adjoint elements span the algebra (Proposition 11.3.10), $p_1 \mathcal{A} = \mathbb{C}\, p_1$. Then

$$\mathcal{A} = (p_1 + \cdots + p_n)\mathcal{A} = \mathbb{C}p_1 + \cdots + \mathbb{C}p_n \simeq \bigoplus_{j=1}^n \mathbb{C}.$$

When $\mathcal{A}$ is not abelian, let $\mathcal{Z}$ be the centre of $\mathcal{A}$. This is an abelian C*-algebra, so the first part of the proof applies. Then we have minimal central projections $p_1, \ldots, p_m$, with $\sum_j p_j = I_\mathcal{A}$, and

$$\mathcal{A} = p_1\mathcal{A} + \cdots + p_m\mathcal{A}. \tag{11.23}$$

Now consider the subalgebra $\mathcal{B} = p_1\mathcal{A}$. If $b$ is a selfadjoint in the centre of $\mathcal{B}$, then $b$ is in the centre of $\mathcal{A}$, since $bp_j = 0$ for $j \geq 2$ and $b$ commutes with all elements in $\mathcal{B}$. This shows that the centre of $\mathcal{B}$ is $\mathbb{C}p_1$, and $p_1$ is the unit of $\mathcal{B}$. As $\dim \mathcal{B} < \infty$, working as in the first paragraph we can find a finite maximal family of pairwise orthogonal minimal subprojections of $p_1$; hence $p_1 = \sum_{j=1}^{n_1} q_j$ with $q_j$ pairwise orthogonal minimal projections. Since each $q_j$ is minimal, $q_j\mathcal{B}q_j = \mathbb{C}q_j$; for if $\dim q_j\mathcal{B}q_j > 1$ Lemma 11.8.3 implies that there exists a proper subprojection of $q_j$.

Fix $q_j$; by Lemma 11.8.5 there exists $c \in \mathcal{B}$ with $q_1 c q_j \neq 0$. Let $q_1 c q_j = ur$ be the Polar Decomposition (Lemma 11.8.9), with $u, r \in \mathcal{B}$. Using that $q_1 c q_j \neq 0$ and (11.22) we have

$$u^*u = R((q_1 c q_j)^*) = R(q_j c^* q_1) = R(q_j c^* q_1 c q_j) \leq q_j,$$

where the last inequality comes from Lemma 11.8.8. Since $q_j$ is minimal, this forces $u^*u = q_j$. Similarly, $uu^* = q_1$. So we have partial isometries $v_{1j} \in \mathcal{B}$, $j = 1, \ldots, n_1$, such that

$$v_{1j}^* v_{1j} = q_j, \quad v_{1j} v_{1j}^* = q_1.$$

Now define $e_{kj} = v_{1k}^* v_{1j}$, $k, j = 1, \ldots, n_1$. These satisfy—because of the orthogonality of the $q_j$ and the equalities $v_{1j} = q_1 1 v_{1j} q_j$—the matrix unit identities:

$$e_{kj} e_{st} = \delta_{j,s}\, e_{kt}, \quad e_{kj}^* = e_{jk}. \tag{11.24}$$

Let $\gamma : M_{n_1}(\mathbb{C}) \to \mathcal{B}$ be given by

$$\gamma\left(\sum_{k,j=1}^{n_1} x_{kj} E_{kj}\right) = \sum_{k,j=1}^{n_1} x_{kj}\, e_{kj}.$$

This map is obviously linear, and the identities (11.24) make it both $*$-preserving and multiplicative. It is injective: if $a = \sum_{k,j=1}^{n_1} x_{kj}\, e_{kj} = 0$, then $0 = e_{1k} a e_{j1} = x_{kj} e_{11}$, so $x_{kj} = 0$ for all $k, j$. It is surjective: if $a \in \mathcal{B}$, then

$$a = I_{\mathcal{A}} a I_{\mathcal{A}} = \sum_{k,j} q_k a q_j = \sum_{k,j} e_{kk} a e_{jj} = \sum_{k,j} e_{k1} e_{11} e_{1k} a e_{j1} e_{11} e_{1j}$$

$$= \sum_{k,j} e_{k1} \alpha_{k,j} e_{11} e_{1j} = \sum_{k,j} \alpha_{k,j}\, e_{kj},$$

where $\alpha_{k,j} \in \mathbb{C}$ is the unique number such that $e_{11}(e_{1k} a e_{j1}) e_{11} = \alpha_{kj} e_{11}$ (using that $e_{11} = q_1$ is minimal). So $\gamma$ is a $*$-isomorphism, and $\mathcal{B} \simeq M_{n_1}(\mathbb{C})$.

Now we can repeat the above for each of $p_j \mathcal{A}$ to obtain $p_j \mathcal{A} \simeq M_{n_j}(\mathbb{C})$, and (11.23) gives

$$\mathcal{A} \simeq \bigoplus_{j=1}^{m} M_{n_j}(\mathbb{C}).$$

□

We conclude the section with more characterizations.

***11.8.11. Proposition.***

*Let $\mathcal{A}$ be a C*-algebra such that every positive element has finite spectrum. Then $\mathcal{A}$ is finite-dimensional.*

*Proof.* Since the spectrum is defined in the unitization of $\mathcal{A}$, we may assume without loss of generality that $\mathcal{A}$ is unital, for showing that the unitization $\tilde{\mathcal{A}}$ is finite-dimensional implies that $\mathcal{A}$ is finite-dimensional.

If every positive element has one-point spectrum, then $\mathcal{A} = \mathbb{C}$; indeed, if $a \in \mathcal{A}^+$ and $\sigma(a) = \{\lambda\}$ then $a - \lambda I_{\mathcal{A}}$ is normal with spectrum $\{0\}$, so

$a - \lambda I_{\mathcal{A}} = 0$ (from $\|a - \lambda I_{\mathcal{A}}\| = \operatorname{spr}(a - \lambda I_{\mathcal{A}})$). It follows that if $\mathcal{A} \neq \mathbb{C}$, there exists $a \in \mathcal{A}^{\mathrm{sa}}$ with spectrum $\{\lambda_1, \ldots, \lambda_m\}$, $m \geq 2$. Since the $\lambda_j$ are isolated, the functions $1_{\{\lambda_j\}}$ are continuous on $\sigma(a)$, and so $p_j = 1_{\{\lambda_j\}}(a) \in \mathcal{A}$. So $\mathcal{A}$ has nontrivial projections. Given any nontrivial projection $p$, either $p\mathcal{A}p = \mathbb{C}p$ or, applying the above, there exists a proper subprojection $q \leq p$.

The previous paragraph allows us to do the following construction. Start with a proper projection $p_1 \in \mathcal{A}$. For each of $p_1\mathcal{A}p_1$ and $(I_{\mathcal{A}} - p_1)\mathcal{A}(I_{\mathcal{A}} - p_1)$ the above applies, and either they are one-dimensional or they contain proper projections. If this process continues indefinitely, we end up with infinitely many pairwise orthogonal projections $\{p_n\}$; this is a contradiction because then $a = \sum_n 2^{-n}p_n \in \mathcal{A}$ and it has infinite spectrum $\sigma(a) = \{0\} \cup \{2^{-n} : n\}$. It follows that there exist $p_1, \ldots, p_m$ such that $p_1 + \cdots + p_m = I_{\mathcal{A}}$ and $p_j\mathcal{A}p_j = \mathbb{C}p_j$ for all $j$.

For $k \neq j$, if $p_k\mathcal{A}p_j \neq \{0\}$, let $a \in p_k\mathcal{A}p_j$ be nonzero. Then $a^*a \in p_j\mathcal{A}p_j$, $aa^* \in p_k\mathcal{A}p_k$, both nonzero. As these algebras are one-dimensional, we have $a^*a = \alpha p_j$, $aa^* = \beta p_k$, for $\alpha, \beta \in \mathbb{C}$. From $a^*a \geq 0$ we get $\alpha > 0$, and similarly $\beta > 0$. Then $\alpha = \|a^*a\| = \|aa^*\| = \beta$. Let $b = a/\alpha^{1/2}$. Then $b^*b = p_j$, $bb^* = p_k$. Let $\phi : p_j\mathcal{A}p_j \to p_k\mathcal{A}p_j$ be given by $\phi(x) = bx$, and $\psi : p_k\mathcal{A}p_j \to p_j\mathcal{A}p_j$ be $\psi(x) = b^*x$. These maps are linear, and

$$\psi(\phi(p_jxp_j)) = b^*bp_jxp_j = p_jxp_j,$$

and similarly $\phi(\psi(y)) = y$; thus they are inverses of each other. They map into the right place since $b^*p_k = p_jb^*$ and $bp_j = p_kb$. So $p_j\mathcal{A}p_j$ and $p_k\mathcal{A}p_j$ either have the same dimension (one) or $p_k\mathcal{A}p_j = \{0\}$.

Finally, $\mathcal{A} = \sum_{k,j=1}^m p_j\mathcal{A}p_k$, a sum of finite-dimensional spaces. □

### *11.8.12. Proposition.*

*Let $\mathcal{A}$ be a $C^*$-algebra. The following statements are equivalent:*

*(i) $\mathcal{A}$ is finite-dimensional;*

*(ii) every selfadjoint element of $\mathcal{A}$ has finite spectrum;*

*(iii) every positive element of $\mathcal{A}$ has finite-spectrum;*

*(iv) every hereditary subalgebra of $\mathcal{A}$ has a unit;*

*(v) $\mathcal{A}$ is reflexive as a Banach space;*

*Proof.* (i) $\Longrightarrow$ (ii) Using Theorem 11.8.10 and basic linear algebra, the spectrum of every element of $\mathcal{A}$ is determined by the characteristic polynomial, so it is finite.

(ii) $\Longrightarrow$ (iii) Trivial.

(iii) $\Longrightarrow$ (i) This follows from Proposition 11.8.11.

(i) $\implies$ (v) By Exercise 7.2.9 every finite-dimensional Banach space is reflexive.

(v) $\implies$ (iv) Let $\mathcal{A}_0 \subset \mathcal{A}$ be a C*-subalgebra. Since $\mathcal{A}$ is reflexive, $\mathcal{A}_0$ is reflexive by Proposition 7.2.16. Then $\overline{B_1^{\mathcal{A}_0}(0)}$ is weak-closed by Banach–Alaoglu (Theorem 7.2.13). Then Krein–Milman (Theorem 7.5.11) applies, so $\overline{B_1^{\mathcal{A}_0}(0)}$ has extreme points; and then Proposition 11.3.12 implies that $\mathcal{A}_0$ is unital.

(iv) $\implies$ (i) If $a \in \mathcal{A}^+$, let $\mathcal{A}_0 = \overline{a\mathcal{A}a}$, which is hereditary by Exercise 11.5.10. By hypothesis, $\mathcal{A}_0$ has a unit $p$. We clearly have $\mathcal{A}_0 \subset p\mathcal{A}p$. If $b \in (p\mathcal{A}p)^+$ then there exists $c \in \mathcal{A}$ such that $b = (pcp)^*pcp = pc^*pcpc \leq \|c\|^2\, p$; since $p \in \mathcal{A}_0$ which is hereditary, then $b \in \mathcal{A}_0$ and as the positives span the C*-algebra we get that $p\mathcal{A}p \subset \mathcal{A}_0$ and hence $\mathcal{A}_0 = p\mathcal{A}p$. Note that $a \in \mathcal{A}_0$, since by functional calculus

$$a = \lim_{n\to\infty} ap\Big(\frac{1}{n}\, I_{\tilde{\mathcal{A}}} + a\Big)^{-1} pa \in \overline{a\mathcal{A}a} = \mathcal{A}_0.$$

Then $a$ is invertible in $\mathcal{A}_0$ by Proposition 11.5.23.

By hypothesis we have that $\mathcal{A}$ is unital. If $\mathcal{A}$ has no minimal projection, we would be able to construct a strictly decreasing sequence of projections $I_{\mathcal{A}} = q_0 \geq q_1 \geq \cdots$, and then then the projections $p_k = q_k - q_{k+1}$ are pairwise orthogonal. This allows us to form $a = \sum_k 2^{-k} p_k \in \mathcal{A}$. By the previous paragraph $\mathcal{A}_0 = \overline{a\mathcal{A}a}$ has a unit $p$ and $a$ is invertible in $\mathcal{A}_0$. But $2^{-k} \in \sigma(a)$ for all $k$ implying that $0 \in \overline{\sigma(a)} = \sigma(a)$, a contradiction. This shows that $\mathcal{A}$ has a minimal projection $p_1$. Working on the hereditary subalgebra $(I_{\mathcal{A}} - p_1)\mathcal{A}(I_{\mathcal{A}} - p_1)$, we get another minimal projection $p_2$. If this process produces infinitely many pairwise orthogonal minimal projections, again we can construct $a$ as before. It follows that there exist finitely many pairwise orthogonal minimal projections $p_1, \ldots, p_n$ with $\sum_k p_k = I_{\mathcal{A}}$. Now we want to show that $p_k\mathcal{A}p_k$ is one-dimensional for all $k$. This is not true in general, but here is where we have to use our hypothesis. If $p_k\mathcal{A}p_k$ is not one dimensional, as in the proof of Proposition 11.8.11 there would be an element with spectrum with at least two points; then functional calculus allows us to construct $b_1, b_2 \in p_k\mathcal{A}p_k$ nonzero, positive, with $b_1b_2 = 0$. We cannot get projections right away as in Proposition 11.8.11 because we don't know a priori that the spectrum is discrete. But now out hypothesis tells us that $\overline{b_1\mathcal{A}b_1}$ is unital, and so we have a nonzero projection $q \leq p_k$ with $qb_2 = 0$, so $q \neq p_k$; this contradicts the minimality of $p_k$. Now we are in the same situation as in the proof of Proposition 11.8.11, to conclude that $\dim \mathcal{A} < \infty$. ∎

**11.8.1. Exercises.**

(11.8.1) Let $\mathcal{A}$ be a C*-algebra and $p \in \mathcal{A}$ a projection. Show that $p$ is positive and that $0 \le p \le I_{\mathcal{A}}$.

(11.8.2) Let $\mathcal{A}$ be a C*-algebra and $p_1, \ldots, p_m \in \mathcal{A}$ projections. Show that the following statements are equivalent:

(i) $p_1, \ldots, p_m$ are pairwise orthogonal;

(ii) $\displaystyle\sum_{k=1}^{m} p_k \le I_{\mathcal{A}}$.

(11.8.3) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ nonzero, and $p_1, \ldots, p_n \in \mathcal{A}$ pairwise orthogonal projections. Let

$$a = \sum_{j=1}^{n} \lambda_j p_j.$$

Show that $a$ is normal, $\sigma(a) = \{\lambda_1, \ldots, \lambda_n\}$ when $\sum_j p_j = I_{\mathcal{A}}$ and $\sigma(a) = \{0\} \cup \{\lambda_1, \ldots, \lambda_n\}$ when $\sum_j p_j \neq I_{\mathcal{A}}$; and for any function $f : \mathbb{C} \to \mathbb{C}$,

$$f(a) = \sum_{j=1}^{n} f(\lambda_j) p_j.$$

(11.8.4) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ nonzero, and $p_1, \ldots, p_n \in \mathcal{A}$ pairwise orthogonal projections. Let

$$a = \sum_{j=1}^{n} \lambda_j p_j.$$

Show that if $a$ is a projection, then $\lambda_1 = \cdots = \lambda_n = 1$.

(11.8.5) Let $\mathcal{A}$ be a C*-algebra, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ distinct and nonzero, and also $\mu_1, \ldots, \mu_m \in \mathbb{C}$ distinct and nonzero. Consider projections $p_1, \ldots, p_n$, and $q_1, \ldots, q_m$ in $\mathcal{A}$ with $p_k p_j = q_k q_j = 0$ when $k \neq j$. Show that if

$$\sum_{k=1}^{n} \lambda_k p_k = \sum_{j=1}^{m} \mu_j q_j$$

then $n = m$ and $\lambda_j = \mu_j$, $p_j = q_j$ for all $j$.

CHAPTER

# 12

# Bounded Operators on a Hilbert Space: Part II

With more tools at hand, we continue our study of bounded operators on a Hilbert space. This leads naturally to von Neumann algebras, via the Double Commutant Theorem (12.3.5).

## 12.1. Locally Convex Topologies on $\mathcal{B}(\mathcal{H})$

We have discussed locally convex topologies in Section 5.4 and Chapter 7.

**12.1.1. The sot and the wot.**

**12.1.1. Definition.**

*Let $\mathcal{H}$ be a Hilbert space. The* ***strong operator topology*** *(**sot**) on $\mathcal{B}(\mathcal{H})$ is the locally convex topology induced by the seminorms*

$$p_\xi(T) = \|T\xi\|, \qquad \xi \in \mathcal{H}.$$

*The* ***weak operator topology*** *(**wot**) on $\mathcal{B}(\mathcal{H})$ is the locally convex topology induced by the seminorms*

$$q_{\xi,\eta}(T) = |\langle T\xi, \eta\rangle|, \qquad \xi, \eta \in \mathcal{H}.$$

What one uses in practice is that $T_j \xrightarrow{\text{sot}} T$ means that $T_j\xi \to T\xi$ for all $\xi \in \mathcal{H}$, and $T_j \xrightarrow{\text{wot}} T$ means that $\langle T_j\xi, \eta\rangle \to \langle T\xi, \eta\rangle$ for all $\xi, \eta \in \mathcal{H}$. The sot is precisely the topology of pointwise convergence of operators, while the wot is the topology of weak pointwise convergence of operators.

It is trivial to check that the weak operator topology is weaker than the strong operator topology. While it is possible to discuss completeness for locally convex spaces the issue is technical and we will not address it. In practical terms, one can show that bounded Cauchy nets in either sot or wot are convergent in $\mathcal{B}(\mathcal{H})$ (Exercise 12.1.2); in particular,

$$\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{sot}} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}} \tag{12.1}$$

(Exercise 12.1.4); this holds for balls, but it does not hold in general (see Corollary 12.1.3 and the comment below it).

As with the weak and weak* topologies in Chapter 7, one can show that these topologies are linear and a base of neighbourhoods of 0 for each is given by

$$V_{\xi_1,\ldots,\xi_n} = \{T \in \mathcal{B}(\mathcal{H}) : \ \|T\xi_j\| < 1, \ j = 1, \ldots, n\}$$

and

$$W_{\xi_1,\ldots,\xi_n,\eta_1,\ldots,\eta_n} = \{T \in \mathcal{B}(\mathcal{H}) : \ |\langle T\xi_j, \eta_j\rangle| < 1, \ j = 1, \ldots, n\}$$

respectively. As in the case with the weak and weak*-topologies, the neighbourhoods are often expressed with an $\varepsilon$ instead of 1 for the inequality, but one ends up getting the same family of neighbourhoods (this was mentioned in Remark 7.1.7).

Because the sot is weaker than the norm topology, any sot-continuous linear functional on $\mathcal{B}(\mathcal{H})$ is automatically norm-continuous (that is, bounded). It is not true in general that a norm-continuous functional is sot-continuous. In fact, sot/wot-continuous linear functionals can be explicitly characterized:

***12.1.2. Proposition.***

*Let $\varphi : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ be linear. The following statements are equivalent:*

*(i) $\varphi$ is wot-continuous;*

*(ii) $\varphi$ is sot-continuous;*

*(iii) there exist $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in \mathcal{H}$ such that*

$$\varphi(T) = \sum_{k=1}^{n} \langle T\xi_k, \eta_k\rangle, \qquad T \in \mathcal{B}(\mathcal{H})$$

*Proof.* (i) $\Longrightarrow$ (ii) this is trivial, since sot is stronger than wot.

(ii) $\Longrightarrow$ (iii) By definition of continuity at 0, there exists a sot neighbourhood $V = V_{\xi_1,\ldots,\xi_n}$ of 0 such that $|\varphi(T)| < 1$ for all $T \in V$. If $T\xi_k = 0$ for all $k$ then $cT \in V$ for all $c > 0$, and hence $|\varphi(cT)| < 1$ for all $c > 0$, implying that $\varphi(T) = 0$. When $T\xi_k \neq 0$ for some $k$,

$$\left|\varphi\left(\frac{T}{2(\sum_{k=1}^n \|T\xi_k\|^2)^{1/2}}\right)\right| < 1$$

(since the coefficient is chosen so that the operator is in $V$), and so

$$|\varphi(T)| < 2\Big(\sum_{k=1}^n \|T\xi_k\|^2\Big)^{1/2}. \tag{12.2}$$

As this inequality also holds when $\varphi(T) = 0$, it is valid for all $T$. Let

$$\mathcal{K} = \overline{\operatorname{span}}^{\|\cdot\|} \{(T\xi_1, \ldots, T\xi_n)^\top : T \in \mathcal{B}(\mathcal{H})\} \subset \mathcal{H}^n.$$

Then $\mathcal{K}$ is a Hilbert space and (12.2) shows that the linear functional

$$(T\xi_1, \ldots, T\xi_n)^\top \longmapsto \varphi(T)$$

is well-defined and bounded on $\mathcal{K}$. By the Riesz Representation Theorem there exists $(\eta_1, \ldots, \eta_n)^\top \in \mathcal{H}^n$ such that

$$\varphi(T) = \langle (T\xi_1, \ldots, T\xi_n)^\top, (\eta_1, \ldots, \eta_n)^\top \rangle = \sum_{k=1}^n \langle T\xi_k, \eta_k \rangle.$$

(iii) $\Longrightarrow$ (i) Straightforward. □

This situation of having two topologies that induce the exact same family of continuous linear functionals already arose in Section 7.1. Concretely, in Theorem 7.1.16 we saw that, since the norm and weak-continuous linear functionals on a locally convex space are the same, the closure of convex sets is the same for both topologies. We have the same situation here:

***12.1.3. Corollary.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be convex. Then $\overline{\mathcal{M}}^{\mathrm{sot}} = \overline{\mathcal{M}}^{\mathrm{wot}}$.*

*Proof.* Since the wot topology is weaker than the sot, we have $\overline{\mathcal{M}}^{\mathrm{sot}} \subset \overline{\mathcal{M}}^{\mathrm{wot}}$ and so we may assume without loss of generality that $\mathcal{M}$ is sot-closed. Let $T \in \overline{\mathcal{M}}^{\mathrm{wot}} \setminus \mathcal{M}$. By Theorem 5.7.18, applied in the locally convex space $(\mathcal{B}(\mathcal{H}), \mathrm{sot})$, there exists a sot-continuous functional $\varphi$ and $c > 0$ with

$$\operatorname{Re}\varphi(T) < c < \operatorname{Re}\varphi(S), \qquad S \in \mathcal{M}.$$

Since $T \in \overline{\mathcal{M}}^{\mathrm{wot}}$, there exists a net $\{T_j\} \subset \mathcal{M}$ with $T_j \xrightarrow{\mathrm{wot}} T$. By Proposition 12.1.2, $\varphi$ is wot-continuous. But then

$$\operatorname{Re}\varphi(T) = \lim_j \operatorname{Re}\varphi(T_j) \geq c,$$

a contradiction. □

If we take $\mathcal{M} = \mathcal{K}(\mathcal{H})$, then $\overline{\mathcal{M}} = \mathcal{M} \subsetneq \mathcal{B}(\mathcal{H}) = \overline{\mathcal{M}}^{\text{sot}}$. This last equality was proven in Proposition 10.6.2; it will be soon be an easy consequence of the Double Commutant Theorem.

Looking at Proposition 12.1.2 and Corollary 12.1.3 it becomes a fair question whether the sot and the wot are actually different. But they are different (when $\dim \mathcal{H} = \infty$, of course), and quite so.

The most obvious difference between both topologies is that the adjoint is not continuous in the sot, while it is continuous in the wot (Exercise 12.1.1). For this, let $S \in \mathcal{B}(\mathcal{H})$ be the unilateral shift with respect to the orthonormal basis $\{\xi_n\}$. Then, since $S^*S = I_{\mathcal{H}}$,

$$\|S^n\xi\|^2 = \langle (S^n)^*S^n\xi, \xi\rangle = \langle \xi, \xi\rangle = \|\xi\|^2.$$

Meanwhile,

$$\|(S^n)^*\xi\|^2 = \|(S^*)^n\xi\|^2 = \sum_{k>n}^{\infty} |\langle \xi, \xi_n\rangle|^2 \xrightarrow[n\to\infty]{} 0.$$

So the adjoint maps the sequence $\{(S^n)^*\}$, which converges sot to 0, to the sequence $\{S^n\}$, which does not converge to 0, showing that it is not sot-continuous.

For an even more brutal distinction between the sot and wot, we look at limits of projections. Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ be a net of projections, and suppose that $P_j \xrightarrow{\text{sot}} T$. Then $T$ is a projection. Indeed, for any $\xi \in \mathcal{H}$

$$\begin{aligned}
\|(T - T^2)\xi\| &\leq \|(T - P_j)\xi\| + \|(P_j^2 - P_jT)\xi\| + \|(P_jT - T^2)\xi\| \\
&= \|(T - P_j)\xi\| + \|P_j(P_j - T)\xi\| + \|(P_j - T)T\xi\| \\
&\leq \|(T - P_j)\xi\| + \|(P_j - T)\xi\| + \|(P_j - T)T\xi\| \longrightarrow 0.
\end{aligned}$$

On the other hand, any positive operator in the unit ball is a wot-limit of projections. To see this we mimic the fact that for any $t \in [0, 1]$ we can construct a projection

$$\begin{bmatrix} t & (t - t^2)^{1/2} \\ (t - t^2)^{1/2} & 1 - t \end{bmatrix}. \tag{12.3}$$

Indeed,

***12.1.4. Proposition.***

*Let* $\dim \mathcal{H} = \infty$ *and* $T \in \mathcal{B}(\mathcal{H})$ *positive with* $\|T\| \leq 1$*. Then there exists a net of projections* $\{P_j\}$ *with* $P_j \xrightarrow{\text{wot}} T$*.*

*The projections* $P_j$ *can be chosen to be finite-rank.*

*Proof.* By Proposition 10.4.3 we know that $0 \leq T \leq I_{\mathcal{H}}$. Fix a wot-neighbourhood $\mathcal{W}$ of 0, that is we choose $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in \mathcal{H}$ and put

$$\mathcal{W} = \{S \in \mathcal{B}(\mathcal{H}) : |\langle S\xi_k, \eta_k\rangle| < 1\}.$$

Let
$$L = \operatorname{span}\{\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n\}$$
Because $\dim L < \infty$ and $\dim \mathcal{H} = \infty$, there exists a subspace $L' \subset L^\perp$ with $\dim L' = \dim L$. We then have a natural identification of $L + L'$ with $L^2$, and we can consider operators on $L^2$ as $2 \times 2$ block matrices. Under that identification we can copy the idea from (12.3). To make things concrete, let $Q$ be the orthogonal projection onto $L$, and $R$ the orthogonal projection onto $L'$. Let $V : L \to L'$ be a partial isometry with $V^*V = Q$, $VV^* = R$ (which exists because $\dim L' = \dim L$). Let $S = QTQ$. Then $S - S^2 \geq 0$ (by Exercise 10.4.7) and we define
$$P = S + (S - S^2)^{1/2}V^* + V(S - S^2)^{1/2} + (R - VSV^*).$$
Then $P = P^*$ and $P = P^2$, the former equality being trivial and the latter following from just computing $P^2$ (note that $V = RVQ$, $SV = V^*S = SR = RS = 0$). And, since $QPQ = QTQ$,
$$\langle (P - T)\xi_k, \eta_k\rangle = \langle (P - T)Q\xi_k, Q\eta_k\rangle = \langle Q(P - T)Q\xi_k, \eta_k\rangle = 0,$$
showing that $P - T \in \mathcal{W}$, which is $P \in T + \mathcal{W}$. We have shown that for every wot-neighbourhood $\mathcal{W}$ of $T$, there exists a projection $P_{\mathcal{W}} \in T + \mathcal{W}$. That is, $T = \lim_{\text{wot}} P_{\mathcal{W}}$ is a wot-limit of projections.

In the construction above, $Q$ and $R$ are finite-rank projections, and $V$ is a finite-rank operator. Thus the $P_j$ constructed are finite-rank. $\square$

Proposition 12.1.4 shows that the product of operators is not jointly wot-continuous, even on bounded sets: it is enough to take $T \in \mathcal{B}(\mathcal{H})$ with $\|T\| = 1$ and $T^2 \neq T$. Then we get projections $P_j \xrightarrow{\text{wot}} T$, but $P_j^2 = P_j \xrightarrow{\text{wot}} T \neq T^2$.

**12.1.5. Proposition.**

*Let $T \in \mathcal{B}(\mathcal{H})$ with $\|T\| \leq 1$. Then there exists a net of unitaries $\{U_j\}$ with $U_j \xrightarrow{\text{wot}} T$.*

*Proof.* Adapt the argument from Proposition 12.1.4 (Exercise 12.1.12). $\square$

Propositions 12.1.4 and 12.1.5 give us more evidence that the strong operator topology and the weak operator topology are very different, as we already knew. Still, we have the following:

**12.1.6. Proposition.**

*Let $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ be a net of projections. If $P_j \xrightarrow{\text{wot}} P$ and $P$ is a projection, then $P_j \xrightarrow{\text{sot}} P$. Similarly, if $\{U_j\} \subset \mathcal{B}(\mathcal{H})$ is a net of unitaries, $U_j \xrightarrow{\text{wot}} U$, and $U$ is a unitary, then $U_j \xrightarrow{\text{sot}} U$.*

*Proof.* Exercise 12.1.14. □

*12.1.7. Remark.* While Propositions 12.1.4 and 12.1.5 yield nets of projections and unitaries respectively, even a wot-convergent *sequence* of projections can fail to converge to a projection, and the same happens with unitaries (Exercises 12.1.15 and 12.1.16).

One important characteristic of the weak operator topology is that it has an Alaoglu theorem.

**12.1.8. Proposition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be bounded. Then $\overline{\mathcal{M}}^{\text{wot}}$ is compact.*

*Proof.* Since closed subsets and multiples of compact subsets are compact, it is enough if we show that the unit ball $\mathcal{Z}_0 = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}$ is wot-compact. For this, the argument is very similar to the proof of the Banach–Alaoglu Theorem (7.2.13). Let $B = \overline{B_1^{\mathcal{H}}(0)}$; we know, from Banach–Alaoglu and the fact that Hilbert spaces are reflexive, that $B$ is weakly compact. Consider the topological space

$$X = \prod_{\xi \in B} B,$$

where each component is given the weak topology, and $X$ the product topology. By Tychonoff's Theorem (1.8.24), $X$ is compact. Define $\Gamma : \mathcal{Z}_0 \to X$ by $\Gamma(T) = \{T\xi\}_{\xi \in B_1(0)}$. It is clear by construction that $\Gamma$ is injective. By definition of the weak operator topology we have that $T_j \xrightarrow{\text{wot}} T$ if and only $T_j\xi \xrightarrow{\text{weak}} T\xi$ for all $\xi \in B$. As convergence in $X$ is pointwise, this means that $\Gamma$ is bicontinuous, and hence a homeomorphism. So the theorem will be proved if we show that $\Gamma\big(\mathcal{Z}_0\big)$ is closed in $X$.

To show that $\Gamma\big(\mathcal{Z}_0\big)$ is closed, we will use Proposition 1.8.6. Let $\{\Gamma(T_j)\}$ be a net in $\Gamma\big(\mathcal{Z}_0\big)$ such that $\Gamma(T_j) \to S \in X$. This means that $T_j\xi \xrightarrow{\text{weak}} S_\xi \in \mathcal{H}$ for all $\xi \in B$. Define $T : \mathcal{H} \to \mathcal{H}$ by $T\xi = S_\xi$. Then $T_j \xrightarrow{\text{wot}} T$, which gives us that $T$ is linear and also

$$\langle T\xi, \eta\rangle = \lim_j \langle T_j\xi, \eta\rangle \leq \|\xi\| \, \|\eta\|$$

for all $\xi, \eta$, which implies that $\|T\xi\| \leq \|\xi\|$, so $\|T\| \leq 1$. Thus $S = \Gamma(T)$ with $T \in \mathcal{Z}_0$, and so $\Gamma(B)$ is closed and hence compact. $\square$

A natural question in light of Proposition 12.1.8 is what happens with the sot. Is the sot-closure of a bounded set compact? The answer is no in general. Fix $\mathcal{M}$ sot-closed and bounded. Then the identity map $\iota : (\mathcal{M}, \text{sot}) \to (\mathcal{M}, \text{wot})$ is continuous (because wot is weaker than sot) and bijective. If $\mathcal{M}$ is sot-compact, then by Exercise 1.8.38 we get that $\iota$ is a homeomorphism, and then sot and wot agree on $\mathcal{M}$. But we know from Proposition 12.1.4 and the paragraph before it that sot and wot do not agree on $\overline{B_1^{\mathcal{B}(\mathcal{H})}}(0)$, and therefore $\overline{B_1^{\mathcal{B}(\mathcal{H})}}(0)$ is not sot-compact.

Here is yet another case where sot and wot convergence agree:

***12.1.9. Lemma.***

*Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be a net of selfadjoint operators such that $T_j \xrightarrow{\text{wot}} T$ and $T_j \leq T$ for all $j$. Then $T_j \xrightarrow{\text{sot}} T$.*

*Proof.* We will need the following estimate: via the C$^*$-identity and since $T - T_j \geq 0$,

$$\|(T - T_j)^{1/2}\|^2 = \|T - T_j\| \leq \|T\| + \|T_j\| \leq 2\|T\|.$$

We also have that $\langle (T - T_j)\xi, \xi\rangle \to 0$ for all $\xi \in \mathcal{H}$. Then $\|(T - T_j)^{1/2}\xi\|^2 = \langle (T - T_j)\xi, \xi\rangle \to 0$, so

$$\|(T - T_j)\xi\|^2 \leq \|(T - T_j)^{1/2}\|^2 \, \|(T - T_j)^{1/2}\xi\|^2 \leq 2\|T\| \, \|(T - T_j)^{1/2}\xi\|^2 \to 0,$$

meaning that $(T - T_j) \xrightarrow{\text{sot}} 0$. $\square$

This next result will play a crucial role in our dealings with von Neumann algebras.

***12.1.10. Proposition.***

*Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be an increasing net of selfadjoint operators. If $\{T_j\}$ is bounded, then there exists $T \in \mathcal{B}(\mathcal{H})$ such that $T_j \xrightarrow{\text{sot}} T$ and $T = \sup_j T_j$.*

*Proof.* By hypothesis, $\|T_j\| < c$ for all $j$. Fix $\xi \in \mathcal{H}$. Then the net $\{\langle T_j\xi, \xi\rangle\}$ is monotone and bounded above in $\mathbb{R}$, since $\langle T_j\xi, \xi\rangle \leq c\|\xi\|^2$. So there exists a unique number $R_\xi = \lim_j \langle T_j\xi, \xi\rangle$. Using Polarization, we get that $R(\xi, \eta) = \lim_j \langle T_j\xi, \eta\rangle$ exists for all $\xi, \eta$. The linearity of the limit and the sequilinearity of the inner product imply that $R$ is a sesquilinear form. Also, since $|\langle T_j\xi, \eta\rangle| \leq c\|\xi\|\,\|\eta\|$ for all $\xi, \eta \in \mathcal{H}$, the form is bounded. By Proposition 10.1.5 there exists $T \in \mathcal{B}(\mathcal{H})$ with

$$\langle T\xi, \eta\rangle = R(\xi, \eta) = \lim_j \langle T_j\xi\, \eta\rangle.$$

By the monotonicity we also get $T_j \leq T$ for all $j$. Then Lemma 12.1.9 gives us $T_j \xrightarrow{\text{sot}} T$.

Now the supremum. By definition, the supremum is the least upper bound. It's clear from the argument above that $\langle T_j\xi, \xi\rangle \leq \langle T\xi, \xi\rangle$ for all $\xi$, so $T_j \leq T$ and hence $T$ is an upper bound for the net $\{T_j\}$. Now suppose that $T_j \leq S$ for all $j$. Again by the argument above, we get that $\langle T\xi, \xi\rangle = \lim_j \langle T_j\xi, \xi\rangle \leq \langle S\xi, \xi\rangle$, for all $\xi \in \mathcal{H}$. Thus $T \leq S$ and hence $T$ is the least upper bound of the net $\{T_j\}$. $\square$

*12.1.11. Remark.* While sot-convergent sequences are bounded (Exercise 12.1.5) the same cannot be said of convergent nets in general. Fix $A$ to be an injective trace-class operator, so that $AP \neq 0$ for all projections $P$. For example, we could take $A = \sum_k 2^{-k}\,\langle\,\cdot\,, \xi_k\rangle\,\xi_k$ for a given orthonormal basis.

Let $\mathcal{F} = \{F \subset \mathcal{H} :\ F \text{ is a finite-dimensional subspace}\}$, ordered by inclusion. We construct a net of operators indexed by $\mathcal{F}$ as follows:

$$T_F = \frac{1}{\operatorname{Tr}(AP_{F^\perp})}\, P_{F^\perp},$$

where $P_{F^\perp}$ is the orthogonal projection onto $F^\perp$. Then, for any $\xi \in \mathcal{H}$, if we move far enough along the net we will have $\xi \in F$, so $T_F\xi = 0$, and then $T_F \to 0$ in the sot topology. If we consider the subspace $F = \operatorname{span}\{\xi_1, \ldots, \xi_m\}$, then

$$\operatorname{Tr}(AP_{F^\perp}) = \sum_{k>m} 2^{-k} = 2^{-m}.$$

So $\|T_F\| = 2^m\|P_{F^\perp}\| = 2^m$ and the sot-convergent net $\{T_F\}$ is not bounded.

*12.1.12. Remark.* Using Remark 7.1.13 we can construct an example where multiplication is not jointly continuous in the strong operator topology. Indeed, let $\left\{\sqrt{n_j}\,\xi_{n_j}\right\}$ be a net in $\mathcal{H}$ that converges weakly to 0. Define $T_j \in \mathcal{B}(\mathcal{H})$ by

$$T_j\xi = \sqrt{n_j}\,\langle \xi, \xi_{n_j}\rangle\,\xi_{n_j}.$$

Then $T_j$ is rank-one and positive. And

$$\|T_j\xi\| = |\sqrt{n_j}\,\langle \xi, \xi_{n_j}\rangle| \to 0,$$

so $T_j \xrightarrow{\text{sot}} 0$. We will see, however, that $T_j^2$ does **not** converge sot to 0. It is straightforward to check that

$$T_j^2\xi = n_j\,\langle \xi, \xi_{n_j}\rangle\,\xi_{n_j},$$

and hence $\|T_j^2\xi\| = n_j\,|\langle \xi, \xi_{n_j}\rangle|$. Let $\xi = \sum_n \frac{1}{n}\,\xi_n$. Then $T_j^2\xi = \xi_{n_j}$, so $\|T_j^2\xi\| = 1$ for all $j$ and therefore $T_j^2\xi$ cannot converge to 0.

Multiplication is jointly sot-continuous, though, when the net on the left is bounded:

***12.1.13. Proposition.***

*Let $\{S_j\}, \{T_j\} \subset \mathcal{B}(\mathcal{H})$ be nets with $S_j \xrightarrow{\text{sot}} S$, $T_j \xrightarrow{\text{sot}} T$ and $\|S_j\| < c$ for all $j$. Then $S_jT_j \xrightarrow{\text{sot}} ST$.*

*Proof.* For any $\xi \in \mathcal{H}$,

$$\begin{aligned}\|(S_jT_j - ST)\xi\| &\le \|S_j(T_j - T)\xi\| + \|(S_j - S)T\xi\| \\ &\le c\,\|(T_j - T)\xi\| + \|(S_j - S)T\xi\| \to 0.\end{aligned}$$

□

*12.1.14. Remark.* In general, there is no reason why two nets should have the same index set, but we can still talk about joint convergence.

*12.1.15. Remark* . The restriction in Proposition 12.1.13 to the case where the net on the left is bounded is not a technicality. Indeed, it is possible to have sot-convergent nets $\{S_j\}$ and $\{T_j\}$, with the latter bounded, and such that $S_jT_j$ does not converge sot. Consider the unbounded convergent net from Remark 12.1.12, $S_j\xi = \sqrt{n_j}\,\langle \xi, \xi_{n_j}\rangle\,\xi_{n_j}$. Define another net $\{T_j\}$ by fixing some $j_0$ and letting $T_j\xi = \frac{1}{\sqrt{n_j}}\langle \xi, \xi_{j_0}\rangle \xi_{n_j}$. Then $\|T_j\xi\| = n_j^{-1/2}\,|\langle \xi, \xi_{j_0}\rangle| \to 0$. So $S_j \xrightarrow{\text{sot}} 0$, $T_j \xrightarrow{\text{sot}} 0$, and $\|T_j\| \le 1$ for all $j$. But

$$\|S_jT_j\xi\| = |\langle \xi, \xi_{j_0}\rangle|$$

does not converge to 0. Of course, $T_jS_j \xrightarrow{\text{sot}} 0$ as predicted by Proposition 12.1.13, since $\|T_jS_j\xi\| = |\langle \xi, \xi_{n_j}\rangle| \to 0$.

***12.1.16. Corollary.***

*Let $T \in \mathcal{B}(\mathcal{H})$ with $0 \leq T \leq I_{\mathcal{H}}$. Then $P = \lim_{\text{sot}} T^{1/n}$ exists and it is the range projection of $T$.*

*Proof.* Because $T$ is positive we can apply continuous functional calculus as in Corollary 11.2.6, so $T^{1/n}$ exists. Also, as $t^{1/n} \leq t^{1/(n+1)}$ in $C(\sigma(T))$, again by continuous functional calculus we get $T^{1/n} \leq T^{1/(n+1)}$ for all $n \in \mathbb{N}$. The sequence $\{T^{1/n}\}$ is also bounded, since $\|T^{1/n}\| \leq 1$ for all $n$ (given that $|t^{1/n}| \leq 1$ in $C(\sigma(T))$). By Proposition 12.1.10, $P = \lim_{\text{sot}} T^{1/n} \in \mathcal{B}(\mathcal{H})^+$ exists. By Proposition 12.1.13,

$$P^2 = \lim_{\text{sot}} T^{1/n}T^{1/n} = \lim_{\text{sot}} T^{2/n} = \lim_{\text{sot}} T^{2/2n} = P.$$

So $P$, being also selfadjoint, is a projection. We also have

$$PT = \lim_{\text{sot}} T^{1/n}T = T,$$

since $t\,t^{1/n} \to t$ (in norm!) in $C(\sigma(T))$. Finally, suppose that $\xi \in (T\mathcal{H})^\perp = \ker T$. Then $P\xi = \lim_n T^{1/n}\xi = 0$. Hence $P$ is the orthogonal projection onto $\overline{T\mathcal{H}}$; that is, the range projection of $T$. $\square$

Recall Definition 10.5.8: given projections $P, Q \in \mathcal{B}(\mathcal{H})$, the **intersection** of $P$ and $Q$ is the projection $P \wedge Q$ onto the subspace $P\mathcal{H} \cap Q\mathcal{H}$.

***12.1.17. Proposition.***

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Then $P \wedge Q = \lim_{\text{sot}} (PQ)^n = \lim_{\text{sot}} (PQP)^n$.*

*Proof.* We note first that the sequence $\{(PQP)^n\}$ is monotone decreasing. Indeed, we have $PQP \leq P^2 = P \leq I_{\mathcal{H}}$ and $(PQP)^2 = PQPQP \leq PQ^2P = PQP$. Now, inductively,

$$\begin{aligned}(PQP)^{2n+1} &= (PQP)^n(PQP)(PQP)^n \\ &\leq (PQP)^nP^2(PQP)^n = (PQP)^{2n}\end{aligned}$$

and

$$\begin{aligned}(PQP)^{2n+2} &= (PQP)^n(PQP)^2(PQP)^n \\ &\leq (PQP)^n(PQP)(PQP)^n = (PQP)^{2n+1}.\end{aligned}$$

So the sequence $\{(PQP)^n\}$ is decreasing and bounded (both ways, but in particular below), so by Proposition 12.1.10 it has a selfadjoint sot-limit $R$

(properly, we apply the proposition to the sequence $\{-(PQP)^n\}$). As sot limits are multiplicative on bounded nets (Proposition 12.1.13), $R^2 = R$ and thus $R$ is a projection. We have

$$RP = \lim_{\text{sot}}(PQP)^n P = \lim_{\text{sot}}(PQP)^n = R.$$

Also—using again Proposition 12.1.13—

$$\begin{aligned} R(I_{\mathcal{H}} - Q)R &= \lim_{\text{sot}}(PQP)^n(I_{\mathcal{H}} - Q)(PQP)^n \\ &= \lim_{\text{sot}}(PQP)^{2n} - \lim_{\text{sot}}(PQP)^n Q(PQP)^n \\ &= R - \lim_{\text{sot}}(PQP)^{2n+1} = R - R = 0. \end{aligned}$$

So $0 = R(I_{\mathcal{H}} - Q)R = [(I_{\mathcal{H}} - Q)R]^*(I_{\mathcal{H}} - Q)R$, giving us $(I_{\mathcal{H}} - Q)R = 0$, or $QR = R$. Taking adjoints we obtain $PR = RQ = R$. So $R \leq P$ and $R \leq Q$. Given $\xi \in R\mathcal{H}$, we have $P\xi = PR\xi = R\xi = \xi$, and similarly $Q\xi = \xi$; so $R\mathcal{H} \subset P\mathcal{H} \cap Q\mathcal{H}$, meaning $R \leq P \wedge Q$. Conversely, let $\xi \in P\mathcal{H} \cap Q\mathcal{H}$. As $\xi = P\xi = Q\xi$, we have $(PQP)\xi = \xi$; inductively, $(PQP)^n\xi = \xi$, and taking limit $R\xi = \xi$. So $P\mathcal{H} \cap Q\mathcal{H} \subset R\mathcal{H}$, showing that $P \wedge Q \leq R$ and thus proving that $R = P \wedge Q$.

Finally,

$$\lim_{\text{sot}}(PQ)^n = \lim_{\text{sot}}(PQP)^{n-1}Q = RQ = R. \qquad \square$$

**12.1.2. The $\sigma$-weak topology.** There is a very natural topology in $\mathcal{B}(\mathcal{H})$, that arises from knowing that it is a dual (Theorem 10.7.11):

**12.1.18. Definition.**

*The* ***$\sigma$-weak topology*** *(also called the* ***ultraweak topology*** *or* ***$\sigma$-wot topology****) is the weak$^*$-topology on $\mathcal{B}(\mathcal{H})$ when seen as the dual of $\mathcal{T}(\mathcal{H})$.*

*The* ***$\sigma$-strong topology*** *(also called* ***ultrastrong****, or* ***$\sigma$-sot****) is the locally convex topology given by the seminorms $p_S(T) = \operatorname{Tr}(ST^*T)^{1/2}$, $S \in \mathcal{T}(\mathcal{H})$.*

By definition, $T_j \xrightarrow{\sigma\text{-weak}} T$ if and only if $\operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST)$ for all $S \in \mathcal{T}(\mathcal{H})$. In this language, $T_j \xrightarrow{\text{wot}} T$ if and only if $\operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST)$ for all $S \in \mathcal{F}(\mathcal{H})$ (Exercise 12.1.13). So wot is weaker than $\sigma$-weak. Meanwhile, $T_j \xrightarrow{\sigma\text{-sot}} T$ if and only if

$$\sum_n \|(T_j - T)\xi_n\|^2 \xrightarrow[j]{} 0 \qquad \text{whenever } \sum_n \|\xi_n\|^2 < \infty,$$

and $T_j \xrightarrow{\text{sot}} T$ if and only if $\operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0$ for all $S \in \mathcal{F}(\mathcal{H})$ (Exercise 12.1.9), so sot is weaker than $\sigma$-sot. The $\sigma$-weak topology is weaker than the $\sigma$-sot, which itself is weaker than the norm topology (Exercise 12.1.18).

*12.1.19. Remark*. The net $\{T_F\}$ from Remark 12.1.12 gives an example of a net that converges wot (sot, in fact) to 0 but diverges ultraweakly. Indeed, with the notation of Remark 12.1.12, for any $F$

$$\operatorname{Tr}(AT_F) = \frac{\operatorname{Tr}(AP_{F^\perp})}{\operatorname{Tr}(AP_{F^\perp})} = 1.$$

So $\{T_F\}$ does not converge ultraweakly. Properly, this shows that $T_F$ does not converge to 0 ultraweakly. But it doesn't converge at all. Indeed, if $P_\xi$ is the orthogonal projection onto $\mathbb{C}\xi$, we have

$$\operatorname{Tr}(P_\xi T_F) = \frac{1}{\operatorname{Tr}(AP_{F^\perp})} \langle T_{F^\perp}\xi, \xi\rangle \to 0.$$

If $T_F \xrightarrow{\sigma\text{-weak}} T$, for ultraweak convergence we need to have $\operatorname{Tr}(P_\xi T_F) \to \operatorname{Tr}(P_\xi T)$. This means that $0 = \operatorname{Tr}(P_\xi T) = \langle T\xi, \xi\rangle$ for all $\xi \in \mathcal{H}$, so $T = 0$. But $T$ cannot be 0 as we would contradict $\operatorname{Tr}(AT) = \lim_F \operatorname{Tr}(AT_F) = 1$. So $T$ cannot exist.

Also, one can tweak the example and consider the net $\{|F|\,T_F\}$; we will still have $|F|\,T_F \xrightarrow{\text{sot}} 0$ but now $\operatorname{Tr}(A\,|F|\,T_F) \to \infty$.

The name "$\sigma$-weak" is justified by this next result. We write $\tilde{\mathcal{H}} = \Big(\bigoplus_{n\in\mathbb{N}} \mathcal{H}\Big)_{\ell^2}$.

**12.1.20. Lemma.**

*Let $\rho : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\tilde{\mathcal{H}})$ be given by $\rho(T) = \bigoplus_n T$. Then $\rho$ is a ${}^*$-monomorphism, and given a net $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ we have $T_j \xrightarrow{\sigma\text{-weak}} 0$ if and only if $\rho(T_j) \xrightarrow{\text{wot}} 0$.*

*Proof.* That $\rho$ is a $*$-homomorphism and that it is injective is entirely straightforward. Now suppose that $T_j \xrightarrow{\sigma\text{-weak}} 0$. Let $\tilde{\xi} \in \tilde{\mathcal{H}}$. This means that $\xi = \{\xi_n\}$, with $\sum_n \|\xi_n\|^2 = \|\tilde{\xi}\|^2$. Fix an orthonormal countable set $\{\eta_n\} \subset \mathcal{H}$ and define $S \in \mathcal{B}(\mathcal{H})$ by $S\eta_n = \xi_n$ and $S = 0$ on the orthogonal complement of $\{\eta_n\}$, extended by linearity. We have $S \in \mathcal{B}(\mathcal{H})$ because for $\nu = \sum_{k=1}^m c_k\eta_k$ we have

$$\|S\nu\| = \left\|\sum_{k=1}^m c_k S\eta_k\right\| = \left\|\sum_{k=1}^m c_k \xi_k\right\|$$

$$\begin{aligned} &\leq \sum_{k=1}^{m} |c_k|\, \|\xi_k\| \leq \Big(\sum_{k=1}^{m} |c_k|^2\Big)^{1/2} \Big(\sum_{k=1}^{m} \|\xi_k\|^2\Big)^{1/2} \\ &\leq \|\tilde{\xi}\|\, \|\nu\|. \end{aligned}$$

Also,

$$\operatorname{Tr}(S^*S) = \sum_n \langle S^*S\eta_n, \eta_n\rangle = \sum_n \|S\eta_n\|^2 = \sum_n \|\xi_n\|^2 = \|\tilde{\xi}\|^2,$$

so $S^*S \in \mathcal{T}(\mathcal{H})$. Hence $SS^* \in \mathcal{T}(\mathcal{H})$ by Exercise 10.7.9, and

$$\begin{aligned} \langle \rho(T_j)\tilde{\xi}, \tilde{\xi}\rangle &= \sum_n \langle T_j\xi_n, \xi_n\rangle = \sum_n \langle T_j S\eta_n, S\eta_n\rangle \\ &\leq \operatorname{Tr}(S^*T_jS) = \operatorname{Tr}(SS^*T_j) \to 0. \end{aligned}$$

After using polarization, we have shown that $\rho(T_j) \xrightarrow{\text{wot}} 0$.

Conversely, suppose that $\rho(T_j) \xrightarrow{\text{wot}} 0$. Fix $S \in \mathcal{T}(\mathcal{H})$. By Theorem 10.6.26 there exist orthonormal sequences $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}$ such that we can write

$$S = \sum_{n=1}^{\infty} \sigma_n \,\langle \cdot, \xi_n\rangle\, \eta_n,$$

where $\{\sigma_n\}$ are the singular values of $S$. Then, after extending $\{\xi_n\}$ to an orthonormal basis,

$$\begin{aligned} \operatorname{Tr}(ST_j) = \operatorname{Tr}(T_jS) &= \sum_{n=1}^{\infty} \langle T_jS\xi_n, \xi_n\rangle = \sum_{n=1}^{\infty} \sigma_n \langle T_j\eta_n, \xi_n\rangle \\ &= \langle \rho(T_j)\tilde{\eta}, \tilde{\xi}\rangle \to 0, \end{aligned}$$

where $\tilde{\eta} = \{\sigma_n^{1/2}\eta_n\} \in \tilde{\mathcal{H}}$ and $\tilde{\xi} = \{\sigma_n^{1/2}\xi_n\} \in \tilde{\mathcal{H}}$ by (10.24). So $T_j \xrightarrow{\sigma\text{-weak}} 0$. □

***12.1.21. Lemma.***

*The $\sigma$-weak and the wot topologies agree on bounded sets. The $\sigma$-sot and the sot topologies agree on bounded sets.*

*Proof.* The wot is weaker than the $\sigma$-weak, so we need to show that if $T_j \xrightarrow{\text{wot}} T$ and $\|T_j\| \leq c$ for all $j$, then $T_j \xrightarrow{\sigma\text{-weak}} T$. Fix $\varepsilon > 0$ and let $S \in \mathcal{T}(\mathcal{H})$. By Proposition 10.7.9 there exists $S_0 \in \mathcal{F}(\mathcal{H})$ such that $\|S - S_0\|_1 < \varepsilon$. Then

$$\begin{aligned} |\operatorname{Tr}(S(T_j - T)| &\leq \big|\operatorname{Tr}\big((S - S_0)(T_j - T)\big)| + |\operatorname{Tr}(S_0(T - T_j))| \\ &\leq \|S - S_0\|_1\, \|T_j - T\| + |\operatorname{Tr}(S_0(T - T_j))| \\ &\leq 2c\,\varepsilon + |\operatorname{Tr}(S_0(T - T_j))|. \end{aligned}$$

Then $\limsup_j |\operatorname{Tr}(S(T_j - T))| \leq 2c\varepsilon$ for all $\varepsilon > 0$ and so $\lim_j \operatorname{Tr}(S(T_j - T)) = 0$ by the Limsup Routine. That is, $T_j \xrightarrow{\sigma\text{-weak}} T$.

We leave the sot/$\sigma$-sot part as Exercise 12.1.19. □

### 12.1.3. Exercises.

(12.1.1) Let $\mathcal{H}$ be a Hilbert space. Show that the adjoint map $T \longmapsto T^*$ is wot-continuous on $\mathcal{B}(\mathcal{H})$.

(12.1.2) Show that any bounded sot or wot Cauchy net in $\mathcal{B}(\mathcal{H})$ is convergent.

(12.1.3) Let $\mathcal{H}$ be a separable Hilbert space. Show that the sot and wot are metrizable on the closed unit ball.

(12.1.4) Prove the equalities (12.1):

$$\overline{B_1^{\mathcal{B}(\mathcal{H})}(0)} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{sot}} = \overline{B_1^{\mathcal{B}(\mathcal{H})}(0)}^{\text{wot}}.$$

(12.1.5) Let $\{T_n\}_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ such that $T_n \xrightarrow{\text{sot}} T \in \mathcal{B}(\mathcal{H})$. Show that $\{T_n\}$ is bounded.

(12.1.6) Let $T \in \mathcal{B}(\mathcal{H})$, $\{\xi_j\} \subset \mathcal{H}$ with dense span, and $\{T_\alpha\} \subset \mathcal{B}(\mathcal{H})$ a net such that there exists $c > 0$ with $\|T_\alpha\| \leq c$ for all $\alpha$. Show that if $T_\alpha \xi_j \to T\xi_j$ for al $j$, then $T_\alpha \xrightarrow{\text{sot}} T$.

(12.1.7) Explain why the argument you used in Exercise 12.1.5 does not apply to nets (as guaranteed by Remark 12.1.11).

(12.1.8) Let $K \in \mathcal{K}(\mathcal{H})$ and $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ be a bounded net with $T_j \xrightarrow{\text{sot}} 0$. Show that $\|T_j K\| \to 0$. Show by example that the assertion can fail if the net is unbounded.

(12.1.9) Show that $T_j \xrightarrow{\text{sot}} T$ if and only if $\operatorname{Tr}(S(T_j - T)^*(T_j - T)) \to 0$ for all $S \in \mathcal{F}(\mathcal{H})$.

(12.1.10) Let $\{T_j\} \subset \mathcal{B}(\mathcal{H})$ with $\|T_j\| \leq c$ for all $j$ and families $\{P_j\}$ and $\{Q_j\}$ of pairwise orthogonal projections such that $T_j = Q_j T_j P_j$ for all $j$. Show that $\sum_j T_j$ converges sot.

(12.1.11) Show that in Exercise 12.1.10 it is not possible to relax the hypothesis on the projections to just $T_j = Q_j T_j$ (that is, it is not enough for the ranges to be pairwise orthogonal, the domains should be too).

(12.1.12) Prove Proposition 12.1.5. Use ideas from the proof of Proposition 12.1.4 and Exercise 10.4.13. The idea in Exercise 11.2.3 will be needed, too.

(12.1.13) Show that $T_j \xrightarrow{\text{wot}} T$ if and only if $\operatorname{Tr}(ST_j) \to \operatorname{Tr}(ST)$ for all $S \in \mathcal{F}(\mathcal{H})$.

(12.1.14) Prove Proposition 12.1.6.

(12.1.15) Find an example of a wot-convergent sequence of projections such that its limit is not a projection.

(12.1.16) Find an example of a wot-convergent sequence of unitaries such that its limit is not a unitary.

(12.1.17) Show that the sot-closure of $\mathcal{U}(\mathcal{H})$ is the set of isometries.

(12.1.18) Show that both sot and $\sigma$-weak are weaker than $\sigma$-sot, which is weaker than the norm topology.

(12.1.19) Show that the sot and $\sigma$-sot agree on bounded sets.

(12.1.20) Let $P, Q \in \mathcal{B}(\mathcal{H})$ be projections. Show that

$$I_{\mathcal{H}} - P \wedge Q = (I_{\mathcal{H}} - P) \vee (I_{\mathcal{H}} - Q).$$

(12.1.21) For each $k, j = 1, \ldots, n$ consider a net $\{T_{k,j,\alpha}\}_\alpha \subset \mathcal{B}(\mathcal{H})$. Form the $n \times n$ matrices $\tilde{T}_\alpha = [T_{k,j,\alpha}]$.

(i) Show that $\tilde{T}_\alpha \xrightarrow{\text{wot}} \tilde{T}$ in $\mathcal{B}(\mathcal{H}^n)$ if and only if $T_{k,j,\alpha} \xrightarrow{\text{wot}} T_{k,j}$ for each $k, j$.

(ii) Show that $\tilde{T}_\alpha \xrightarrow{\text{sot}} \tilde{T}$ in $\mathcal{B}(\mathcal{H}^n)$ if and only if $T_{k,j,\alpha} \xrightarrow{\text{sot}} T_{k,j}$ for each $k, j$.

(12.1.22) Let $\{P_j\}_{j\in J} \subset \mathcal{B}(\mathcal{H})$ be a net of pairwise orthogonal projections. Show that the series $P = \sum_j P_j$ converges sot, and $P = \bigvee_j P_j$.

(12.1.23) Let $\mathcal{A}$ be a C*-algebra, $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation, and $\{e_j\}$ an approximate unit for $\mathcal{A}$. Show that $\pi(e_j) \xrightarrow{\text{sot}} I_{\mathcal{H}}$.

(12.1.24) Let $\{Q_j\} \subset \mathcal{B}(\mathcal{H})$ be a non-increasing net of projections. Show that $\bigwedge_j Q_j = \lim_{\text{sot}} Q_j$.

## 12.2. Multiplication Operators

One natural way to construct selfadjoint and normal operators is that of **multiplication operators**.

**12.2.1. Definition.**

*Let $(X, \mu)$ be a measure space, and consider $\mathcal{H} = L^2(X, \mu)$. Given $f \in L^\infty(X, \mu)$ we define the **multiplication operator** $M_f : \mathcal{H} \to \mathcal{H}$ by $M_f g = fg$.*

Here are some basic facts:

***12.2.2. Proposition.***

*Let $(X, \mu)$ be a measure space. Given $f, g \in L^\infty(X)$,* *(i)* $M_f \in \mathcal{B}(L^2(X))$, *and* $\|M_f\| = \|f\|_\infty$*;*

*(ii)* $M_f + M_g = M_{f+g}$ *and* $M_f M_g = M_{fg}$*;*

*(iii)* $M_f^* = M_{\bar f}$*, so* $M_f$ *is always normal, and it is selfadjoint if and only if* $\bar f = f$*, i.e.,* $f$ *is real a.e.;*

*(iv) If* $(X,\mu)$ *is semifinite (in particular, if* $\sigma$*-finite) then* $\sigma(M_f) = \operatorname{ess\,ran} f$*;*

*Proof.* Exercise 12.2.1. $\square$

We will eventually see that, with the right point of view, any normal operator can be seen as a multiplication operator; that is the content of the Spectral Theorem. Meanwhile, let us show some results that will help in said quest.

***12.2.3. Proposition.***

*Let* $(X,\Sigma,\mu)$ *be a* $\sigma$*-finite measure space. Consider* $L^\infty(X,\mu)$ *as the dual of* $L^1(X,\mu)$*. Then a net* $\{g_j\} \subset L^\infty(X,\mu)$ *converges weak* to* $g$ *if and only if* $M_{g_j}$ *converges to* $M_g$ *in the weak operator topology of* $\mathcal{B}(L^2(X,\mu))$*.*

*Proof.* Since the topologies are linear we may assume without loss of generality that $g = 0$. Suppose that $M_{g_j} \xrightarrow{\text{wot}} 0$. Via polarization, this means that $\langle g_j h, h\rangle \to 0$ for all $h \in L^2(X,\mu)$, which we can write as

$$\int_X g_j\,|h|^2\,d\mu \to 0.$$

Given $h \in L^1(X,\mu)$, we can write it as a linear combination of four positive functions, all in $L^1(X,\mu)$. If one of them is $h_1$, say, we have that $h_1^{1/2} \in L^2(X,\mu)$. Therefore

$$\int_X g_j h_1\,d\mu = \langle M_{g_j} h_1^{1/2}, h_1^{1/2}\rangle \to 0.$$

By linearity we get the same limit for $h$. So $g_j \xrightarrow{\text{weak}^*} 0$. And the converse also holds: if $\int_X g_j h\,d\mu \to 0$ for all $h \in L^1(X,\mu)$, then $\langle g_j h, h\rangle \to 0$ for all $h \in L^2(X,\mu)$, since $|h|^2 \in L^1(X,\mu)$. $\square$

The $\sigma$-finiteness in Proposition 12.2.3 can be relaxed a bit, but not a lot because we lose the duality. This was mentioned in Remark 5.6.12.

**12.2.1. Exercises.**

(12.2.1) Prove Proposition 12.2.2.

(12.2.2) Show that the semifinite hypothesis is crucial for (iv) to hold in Proposition 12.2.2.

(12.2.3) Show that $\lambda \in \sigma(M_f)$ is an eigenvalue with multiplicity $m$ if and only if $\{f = \lambda\}$ consists of exactly $m$ atoms.

## 12.3. Commutants and Double Commutants

The notion of commutant plays a central role both in the study of representations, and in the study of von Neumann algebras.

**12.3.1. Definition.**

*Given a set $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, its* ***commutant*** *is the set*

$$\mathcal{M}' = \{S \in \mathcal{B}(\mathcal{H}) : \ ST = TS \text{ for all } T \in \mathcal{M}\}.$$

*The* ***double commutant*** *of $\mathcal{M}$ is*

$$\mathcal{M}'' = (\mathcal{M}')'.$$

First, some rather straightforward consequences of the definition, that will be used repeatedly.

***12.3.2. Proposition.***

*Let $\mathcal{M}, \mathcal{N} \subset \mathcal{B}(\mathcal{H})$. Then*

*(i) For any $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, $\mathbb{C}\, I_{\mathcal{H}} \subset \mathcal{M}'$.*

*(ii) If $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is an algebra, then $\mathcal{A}$ is abelian if and only if $\mathcal{A} \subset \mathcal{A}'$.*

*(iii) $\mathcal{M}'$ is a wot-closed algebra (hence, sot-closed too);*

*(iv) if $\mathcal{N} \subset \mathcal{M}$ then $\mathcal{M}' \subset \mathcal{N}'$;*

*(v) if $\mathcal{M}^* = \mathcal{M}$, then $\mathcal{M}'$ is a $*$-algebra;*

*(vi) $\mathcal{M} \subset \mathcal{M}''$;*

*(vii) $\mathcal{M}' = \mathcal{M}'''$.*

*Proof.* Exercise 12.3.6. □

*12.3.3. Examples.* It is not generally easy to determine the commutant of a set, but here are some basic canonical examples. We mostly care about commutants of $*$-algebras.

(a) $\mathcal{B}(\mathcal{H})' = \mathbb{C}\, I_{\mathcal{H}}$, and $(\mathbb{C}\, I_{\mathcal{H}})' = \mathcal{B}(\mathcal{H})$ (Exercise 12.3.1).

(b) Let $D \in M_n(\mathbb{C})$ be diagonal with all diagonal entries distinct; then $\{D\}' = \{D\}'' = \{E \in M_n(\mathbb{C}) : \text{ diagonal}\}$ (Exercise 12.3.4). We will see a generalization of this in Proposition 12.3.4, and another right below.

(c) Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space and $\{\xi_n\}$ an orthonormal basis. Given $a \in \ell^\infty(\mathbb{N})$, let $M_a \in \mathcal{B}(\mathcal{H})$ be the linear operator induced by $M_a \xi_n = a_n \xi_n$, and $\mathcal{A} = \{M_a : \ a \in \ell^\infty(\mathbb{N})\}$ the algebra of multiplier operators. Then $\mathcal{A} = \mathcal{A}'$ (Exercise 12.3.2). This is often called the **diagonal masa**; properly, it is one of them, as the construction can be made for any orthonormal basis (see Definition 12.5.5 and Proposition 12.5.8).

Given Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$, $n \in \mathbb{N} \cup \{\infty\}$, we consider

$$\mathcal{H}^{(n)} = \left( \bigoplus_{k=1}^{n} \mathcal{H} \right)_{\ell^2}$$

(cfr. Definition 5.3.8) as a Hilbert space with the inner product

$$\langle (\xi_k), (\eta_k) \rangle = \sum_{k=1}^{n} \langle \xi_k, \eta_k \rangle.$$

When $n = \infty$ the condition $\sum_k \|\xi_k\|^2 < \infty$ guarantees that the inner product above is well-defined.

When $n < \infty$, there is a canonical identification between the algebras $\mathcal{B}(\mathcal{H}^{(n)})$ and $M_n(\mathcal{B}(\mathcal{H}))$ where each $T \in \mathcal{B}(\mathcal{H}^n)$ is seen as an $n \times n$ matrix with entries $T_{kj} \in \mathcal{B}(\mathcal{H}_j, \mathcal{H}_k)$,

$$T_{kj}\xi = \pi_k T(\xi \otimes e_j),$$

where $\xi \otimes e_j = (0, \ldots, 0, \xi, 0, \ldots)$, with $\xi$ in the $j^{\text{th}}$ position, and $\pi_k((\xi_h)) = \xi_k$. As in linear algebra, this matrix representation transforms composition of operators into the usual matrix product “row times column”.

**_12.3.4. Proposition._**

*Let* $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$*,* $n \in \mathbb{N} \cup \{\infty\}$*. Let* $\mathcal{M}^{(n)} = \bigoplus_{k=1}^{n} \mathcal{M} \subset \mathcal{B}(\mathcal{H}^n)$*. Then*

$$(\mathcal{M}^{(n)})' = \{T \in \mathcal{B}(\mathcal{H}^{(n)}) : T_{kj} X_j = X_k T_{kj} \ \ \textit{for all } X \in \mathcal{M}^{(n)}, \ k, j\},$$

*and*

$$(\mathcal{M}^{(n)})'' = \{T \in \mathcal{B}(\mathcal{H}^{(n)}) : T_{kk} \in \mathcal{M}'', \ T_{kj} = 0 \ \ \textit{if } k \neq j\}.$$

*Proof.* Exercise 12.3.7. □

We now come into one of the first of many Murray–von Neumann gems:

**_12.3.5. Theorem. (Double Commutant)_**

*Let* $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ *be a non-degenerate selfadjoint set. Then*

$$\mathcal{M}'' = \overline{\operatorname{alg}(\mathcal{M})}^{\mathrm{sot}}.$$

*In particular, if* $\mathcal{M}$ *is a* $*$*-algebra then* $\mathcal{M}'' = \overline{\mathcal{M}}^{\mathrm{sot}}$*.*

*Proof.* As $\mathcal{M}$ is selfadjoint we may assume without loss of generality that it is a $*$-algebra. Since $\mathcal{M}''$ is sot-closed and $\mathcal{M} \subset \mathcal{M}''$, then $\overline{\mathcal{M}}^{\mathrm{sot}} \subset \mathcal{M}''$.

Now suppose that $T \in \mathcal{M}''$. The assumption that $\mathcal{M}$ is non-degenerate means that $\mathcal{H} = \overline{\mathcal{M}\mathcal{H}}$. Fix $\xi \in \mathcal{H}$ and let $\mathcal{H}_0 = \overline{\mathcal{M}\xi}$. Since this is a closed subspace of $\mathcal{H}$, there exists an orthogonal projection $P$ onto $\mathcal{H}_0$ (Proposition 4.3.8). For any $S, R \in \mathcal{M}$,

$$PS\,R\xi = P\,(SR)\xi = SR\xi = S\,R\xi.$$

So $PS = S$ on $\mathcal{M}\xi$, and by taking limits $PS = S$ on $\mathcal{H}_0$. That is, $PSP = SP$. For fixed $S$ we can do the same for $S^*$, and so $PS^*P = S^*P$; taking adjoints, $PSP = PS$. Comparing with the previous expression, $SP = PS$. That is, $P \in \mathcal{M}'$. In particular, $TP = PT$. Now, since $PT\xi \in \mathcal{H}_0$, there exists a sequence $\{T_k\} \subset \mathcal{M}$ with

$$T\xi = TP\xi = PT\xi = \lim_k T_k\xi.$$

In particular, given $\varepsilon > 0$ there exists $T_0 \in \mathcal{M}$ with $\|(T - T_0)\xi\| < \varepsilon$. Now fix $\varepsilon > 0$ and $\xi_1, \dots, \xi_n \in \mathcal{H}$. Consider the $*$-algebra $\tilde{\mathcal{M}} = \{(T, \dots, T)^\top : T \in \mathcal{M}\} \subset \mathcal{B}(\mathcal{H}^n)$, and $\tilde{\xi} = \bigoplus_{j=1}^{n} \xi_j \in \mathcal{H}^{(n)}$. By Proposition 12.3.4,

$$\tilde{T} = \bigoplus_j T \in \bigoplus_{j=1}^{n} \mathcal{M}''.$$

By the first part of the proof, there exists $\tilde{T}_0 \in \tilde{\mathcal{M}}$ such that $\|(\tilde{T} - \tilde{T}_0)\tilde{\xi}\| < \varepsilon$. Unpacked, this is

$$\sum_{j=1}^{n} \|(T - T_0)\xi_j\|^2 < \varepsilon^2.$$

In particular, $\|(T - T_0)\xi_j\| < \varepsilon$, $j = 1, \ldots, n$. We have proven that given any basic sot-neighbourhood $V$ of $T$, that is $V = \{S : \|(T - S)\xi_j\| < \varepsilon,\ j = 1, \ldots, n\}$, there exists $T_0 \in \mathcal{M}$ with $T_0 \in V$. Thus $T \in \overline{\mathcal{M}}^{\text{sot}}$. □

The Double Commutant Theorem is a remarkable result, in that taking double commutants is an algebraic process, while taking sot-closure is a topological one.

We will discuss von Neumann algebras in some detail in Chapter 14, but we will state the definition now, as the terminology will be needed in Section 12.4 and beyond.

**12.3.6. Definition.**

*A $*$-algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is a* **von Neumann algebra** *if $\mathcal{M} = \overline{\mathcal{M}}^{\text{sot}}$ (equivalently, $\mathcal{M} = \overline{\mathcal{M}}^{\text{wot}}$).*

*We denote by $\mathcal{P}(\mathcal{M}) = \{P \in \mathcal{M} : P^*P = P\}$ the set of projections in $\mathcal{M}$.*

*When the centre $\mathcal{Z}(\mathcal{M})$ of $\mathcal{M}$ is trivial (that is, equal to $\mathbb{C}\, I_{\mathcal{M}}$), we say that $\mathcal{M}$ is a* **factor**.

Since the sot and wot topologies are weaker than the norm topology, every von Neumann algebra is a C$^*$-algebra. Combining this with the Double Commutant Theorem, we can say that non-degenerate von Neumann algebras are those C$^*$-algebras $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ such that $\mathcal{M} = \mathcal{M}''$. It is important to notice that, as opposed to C$^*$-algebras, von Neumann algebras are defined as subalgebras of $\mathcal{B}(\mathcal{H})$; so every time we say "let $\mathcal{M}$ be a von Neumann algebra" we are actually saying "let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra". It is possible to characterize which C$^*$-algebras admit a faithful representation as a von Neumann algebra (Theorem 12.7.5), so classically authors would distinguish a *$W^*$-algebra* (meaning a C$^*$-algebra that admits a faithful representation as a von Neumann algebra) and a *von Neumann algebra* (wot/sot-closed C$^*$-subalgebra of $\mathcal{B}(\mathcal{H})$); in practice it is very rare to consider an abstract von Neumann algebra, so the latter terminology has stuck.

The centre $\mathcal{Z}(\mathcal{M})$ is always an abelian von Neumann algebra. It is straightforward to check that $\mathcal{Z}(\mathcal{M}') = \mathcal{M} \cap \mathcal{M}' = \mathcal{Z}(\mathcal{M})$; in particular $\mathcal{M}$ is a factor if and only if $\mathcal{M}'$ is (Exercise 12.3.3).

*12.3.7. Examples.* Below are some very elementary examples of von Neumann algebras.

- Every finite-dimensional C$^*$-algebra is a von Neumann algebra (Exercise 12.3.5).
- $\mathcal{B}(\mathcal{H})$ is a von Neumann algebra.
- $\mathcal{K}(\mathcal{H})$, when $\dim \mathcal{H} = \infty$, is a C$^*$-algebra that is not a von Neumann algebra, since $\overline{\mathcal{K}(\mathcal{H})}^{\text{sot}} = \mathcal{B}(\mathcal{H})$.
- $\ell^\infty(\mathbb{N}) \subset \mathcal{B}(\ell^2(\mathbb{N}))$, seen as multiplication operators, is a von Neumann algebra. This will be very easy to see when some machinery is available.
- More generally, $L^\infty(\mu) \subset \mathcal{B}(L^2(\mu))$, seen as multiplication operators, is a von Neumann algebra whenever $\mu$ is localizable (in particular, if $\sigma$-finite).

It should be emphasized that the above examples are not representative of the full zoo of von Neumann algebras. These come in many flavours, with many of them very different from the examples above. We will see the first non-trivial instance of a von Neumann algebra in Example 13.3.13.

The word "factor" to designate those von Neumann algebras with trivial centre comes from the fact that any von Neumann algebra can be decomposed in terms of them. This is trivial in finite dimension, where any von Neumann algebra is of the form $\bigoplus_j M_{n(j)}(\mathbb{C})$ (Theorem 11.8.10), with each full matrix block being a factor. For an arbitrary von Neumann algebra something similar can be done if one instead of a direct sum uses a **direct integral**; this was shown by von Neumann in [vN49] (the paper completed in 1938). This subject is very technical and seldom used in modern literature, so we will skip it. We refer to [KR97b, Chapter 14] for a very detailed account, even if restricted to the case where $\mathcal{H}$ is separable (this restriction is also in effect in von Neumann's original paper, where he claims that most results would be unaffected in the general case).

As with many other structures, an intersection of von Neumann algebras is again a von Neumann algebra. This allows us to talk, given $\mathcal{X} \subset \mathcal{B}(\mathcal{H})$, about the von Neumann algebra generated by $\mathcal{X}$, which we denote by $W^*(\mathcal{X})$. When the set $\mathcal{X}$ is selfadjoint (that is, $T \in \mathcal{X}$ if and only if $T^* \in \mathcal{X}$) and non-degenerate (that is, $\overline{X\mathcal{H}} = \mathcal{H}$) we have that $W^*(\mathcal{X}) = \mathcal{X}''$.

We follow with a few consequences of the Double Commutant Theorem.

As mentioned before Proposition 11.3.11, if $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is a C$^*$-algebra and $T \in \mathcal{A}$ with polar decomposition $T = V|T|$, it might not be the case that $V \in \mathcal{A}$. But things are different when working in a von Neumann algebra:

**12.3.8. Corollary.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $T \in \mathcal{M}$. If $T = V|T|$ is the unique polar decomposition, then $|T| \in \mathcal{M}$ and $V \in \mathcal{M}$. In particular, the range projection of $T$ is in $\mathcal{M}$.*

*Proof.* The fact that $|T| \in \mathcal{M}$ holds in any C*-algebra, since $|T| = (T^*T)^{1/2}$. As for $V$, we know that it is a partial isometry from $\overline{\operatorname{ran}} T^*$ to $\overline{\operatorname{ran}} T$. Let $S \in \mathcal{M}'$. For any $\xi \in \mathcal{H}$,

$$SV|T|\xi = ST\xi = TS\xi = V|T|S\xi = VS|T|\xi.$$

So $VS = SV$ on $\overline{\operatorname{ran}}|T|$. As $S|T| = |T|S$, $\overline{\operatorname{ran}}|T|$ is invariant by $S$ and $S^*$, and so its orthogonal complement is also invariant by $S$; then $VS = 0 = SV$ on $(\operatorname{ran}|T|)^\perp = (\operatorname{ran} T^*)^\perp$. Hence $SV = VS$. As $S$ was arbitrary, $V \in \mathcal{M}'' = \mathcal{M}$.

As for the range projection, by Proposition 10.4.11 it equals $VV^* \in \mathcal{M}$. □

A different proof of Corollary 12.3.8 can be seen in Exercise 12.4.11. In Definition 10.5.8 we defined the union and the intersection of projections.

**12.3.9. Corollary.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\{P_j\} \subset \mathcal{M}$ a family of projections. Then $\bigvee_j P_j \in \mathcal{M}$ and $\bigwedge_j P_j \in \mathcal{M}$.*

*Proof.* Write $Q = \bigwedge_j P_j$. Fix $T \in \mathcal{M}'$ selfadjoint and $\xi \in \mathcal{H}$. We have, for any $j$, $TQ\xi = TP_jQ\xi = P_jTQ\xi$. That is, $TQ\xi \in \bigcap_j P_j\mathcal{H}$. This says that $TQ\xi = QTQ\xi$. As $\xi$ was arbitrary, $TQ = QTQ$. Taking adjoints, $TQ = QT$. As the selfadjoint elements in $\mathcal{M}'$ span all of $\mathcal{M}'$, we have $QT = QT$ for all $T \in \mathcal{M}'$, and therefore $Q \in \mathcal{M}'' = \mathcal{M}$.

Now, using Proposition 10.5.9,

$$\bigvee_j P_j = I_{\mathcal{M}} - \bigwedge_j (I_{\mathcal{M}} - P_j) \in \mathcal{M}.$$ □

Corollary 12.3.9 shows that the set of projections of a von Neumann algebra is a **lattice** (it contains its own joins and meets). It is in fact a **complete lattice** with respect to the usual order of operators in the sense that it contains the supremum and the infimum of any subset (Exercise 12.3.8).

**12.3.10. Corollary.**

*Let $\mathcal{M}$ be a von Neumann algebra. Then $\mathcal{M}$ is unital.*

*Proof.* Let $E = \bigvee\{P : P \in \mathcal{P}(\mathcal{M})\}$. By Corollary 12.3.9, $E \in \mathcal{M}$. By Corollary 12.1.16, for any $T \in \mathcal{M}^+$ its range projection $P_T$ is in $\mathcal{M}$. Then $P_T \leq E$, so $ET = EP_TT = P_TT = T$. Taking adjoints we also get $TE = T$. As the positive elements span $\mathcal{M}$, $E$ is the unit of $\mathcal{M}$. □

We will use the following notation, that will be fully justified in Section 13.1: if $A \in \mathcal{A}$, we denote by $A \otimes I_n \in M_n(\mathcal{A})$ the matrix with constant diagonal $A$ and zeros elsewhere.

**12.3.11. Corollary.**

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate $C^*$-algebra and $n \in \mathbb{N}$. We consider $M_n(\mathcal{A}) \subset \mathcal{B}(\mathcal{H}^n)$. Then*

$$M_n(\mathcal{A})' = \{X \otimes I_n : X \in \mathcal{A}'\}, \qquad \{X \otimes I_n : X \in \mathcal{A}\}' = M_n(\mathcal{A}')$$

*and*

$$M_n(\mathcal{A})'' = M_n(\mathcal{A}'').$$

*Proof.* Exercise 12.3.10. □

**12.3.12. Corollary.**

*Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{J} \subset \mathcal{M}$ a wot-closed ideal. Then there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ such that $\mathcal{J} = Z\mathcal{M}$.*

*Proof.* Let $Z \in \mathcal{J}$ be its unit, which exists by Corollary 12.3.10. We have $Z^2 = Z = Z^*$, so $Z$ is a projection. Let $T \in \mathcal{M}$ be selfadjoint. Since $ZT \in \mathcal{J}$, we have $ZT = ZTZ$. Taking adjoints, $TZ = ZTZ = ZT$. So $Z$ commutes with all selfadjoints in $\mathcal{M}$, and as these span all of $\mathcal{M}$, we get that $Z \in \mathcal{Z}(\mathcal{M})$. □

**12.3.13. Definition.**

*For a subspace $\mathcal{H}_0 \subset \mathcal{H}$, $[\mathcal{H}_0]$ denotes the orthogonal projection onto $\overline{\mathcal{H}_0}$.*

**12.3.14. Corollary.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $\xi \in \mathcal{H}$. Then $[\mathcal{M}'\xi] \in \mathcal{M}$.*

*Proof.* Let $P = [\mathcal{M}'\xi]$. For any $T \in \mathcal{M}'$, $S \in \mathcal{M}'$, we have $TP(S\xi) = TS\xi = PT(S\xi)$. So $TP = PT$ on $\overline{\mathcal{M}'\xi}$. If $\eta \in (\mathcal{M}'\xi)^{\perp}$, then $T\eta \in (\mathcal{M}'\xi)^{\perp}$, for

$$\langle T\eta, S\xi\rangle = \langle \eta, T^*S\xi\rangle = 0.$$

So $PT\eta = 0$. And $TP\eta = 0$ since $P\eta = 0$. Thus $PT = TP$, and hence $P \in \mathcal{M}'' = \mathcal{M}$. □

We will now focus on a remarkable consequence of the Double Commutant Theorem: Kaplansky's renowned density theorem. As Pedersen put it in [Ped79], using the well-known adage,

> *The density theorem is Kaplansky's great gift to mankind [Kap51b]. It can be used every day, and twice on Sundays.*

**12.3.15. Definition.**

*A function $f : \mathbb{R} \to \mathbb{R}$ is* ***strongly continuous*** *on $\mathcal{B}(\mathcal{H})^{\mathrm{sa}}$ if whenever $\{T_j\}$ is a net of selfadjoint operators with $T_j \xrightarrow{\text{sot}} T$, then $f(T_j) \xrightarrow{\text{sot}} f(T)$.*

Not every continuous function is strongly continuous. For instance we saw in Remark 12.1.12 that the function $f(t) = t^2$ is not strongly continuous.

**12.3.16. Lemma.**

*Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and such that $|f(t)| \leq \alpha + \beta|t|$ for some $\alpha, \beta \geq 0$ and all $t \in \mathbb{R}$. Then $f$ is strongly continuous.*

*Proof.* Let

$$\mathcal{S} = \{f : \mathbb{R} \to \mathbb{R},\ \ \text{strongly continuous}\}, \qquad \mathcal{S}^b = \{f \in \mathcal{S},\ \ \text{bounded}\}.$$

The fact that for any selfadjoint $T$ we have $\|f(T)\| \leq \|f\|_\infty$ (since $\|f(T)\| = \sup\{|f(t)| :\ t \in \sigma(T)\}$) guarantees that $\mathcal{S}$ and $\mathcal{S}^b$ are closed under uniform limits (Exercise 12.3.12).

Given $f \in \mathcal{S}^b$ and $g \in \mathcal{S}$, if $T_j \xrightarrow{\text{sot}} T$ then $fg(T_j) \xrightarrow{\text{sot}} fg(T)$ by Proposition 12.1.13. So $\mathcal{S}^b\mathcal{S} \subset \mathcal{S}$ (and, a fortiori, $\mathcal{S}\mathcal{S}^b \subset \mathcal{S}$ by commutativity). One consequence of this is that $\mathcal{S}^b$ is an algebra, hence a C*-algebra. Now let $h(t) = (1+t^2)^{-1}$ and $g(t) = t\,h(t)$. Using the usual algebraic identity

$a^{-1} - b^{-1} = a^{-1}(b-a)b^{-1}$, for $\xi \in \mathcal{H}$

$$\begin{aligned}
\big\|\big(h(T_j) - h(T)\big)\xi\big\| &= \big\|\big((I_{\mathcal{H}} + T_j^2)^{-1}(T^2 - T_j^2)(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&= \big\|\big((I_{\mathcal{H}} + T_j^2)^{-1}\big((T_j(T - T_j) \\
&\qquad + (T - T_j)T\big)(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\le \big\|\big((I_{\mathcal{H}} + T_j^2)^{-1}T_j\big\|\, \big\|(T - T_j)(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\qquad + \big\|\big((I_{\mathcal{H}} + T_j^2)^{-1}\big\|\, \big\|(T - T_j)T(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\le \|g(T_j)\|\, \big\|(T - T_j)(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\qquad + \|h(T_j)\|\, \big\|(T - T_j)T(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\le \|g\|_\infty\, \big\|(T - T_j)(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\qquad + \|h\|_\infty\, \big\|(T - T_j)T(I_{\mathcal{H}} + T^2)^{-1}\big)\xi\big\| \\
&\to 0.
\end{aligned}$$

Thus $h \in \mathcal{S}^b$. As $\frac{1}{s^2}T_j \xrightarrow{\text{sot}} \frac{1}{s^2}T$ if and only if $T_j \xrightarrow{\text{sot}} T$, we get that $h_s(t) = (1 + s^2t^2)^{-1}$ is in $\mathcal{S}^b$ for all $s \in \mathbb{R}$ (the case $s = 0$ is trivial). We also have that the identity function $z(t) = t$ is in $\mathcal{S}$, so $k_s(t) = st\, h_s(t)$ is in $\mathcal{S}$ for all $s \in \mathbb{R}$. The functions $\{h\} \cup \{k_s : \ s \in \mathbb{R}\}$ are in $C_0(\mathbb{R})$. By Exercise 7.4.6 they generate $C_0(\mathbb{R})$ as a C$^*$-algebra. Hence $C_0(\mathbb{R}) \subset \mathcal{S}^b$.

Consider $f$ as in the hypothesis, that is continuous and such that $|f(t)| \le \alpha + \beta|t|$ for all $t$ and fixed $\alpha, \beta \ge 0$. We have $hf \in C_0(\mathbb{R}) \subset \mathcal{S}^b$. Also, $zhf = gf \in \mathcal{S}\mathcal{S}^b \subset \mathcal{S}$; being also bounded, $gf \in \mathcal{S}^b$. Multiplying by $z$, $z^2hf \in \mathcal{S}$. But then

$$f = hf + z^2hf \in \mathcal{S}. \qquad \square$$

The notation we have been using for balls does not sit very well with the next statement. So we will temporarily use, for $X \subset \mathcal{B}(\mathcal{H})$, the notation $(X)_1 = \{x \in X : \ \|x\| \le 1\}$. Also, to be clear, $\mathcal{A}''_{\text{sa}}$ means the selfadjoint part of $\mathcal{A}''$, and similarly with $\mathcal{A}''_+$.

***12.3.17. Theorem. (Kaplansky's Density Theorem)***

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate $*$-algebra. Then*

- $\overline{(\mathcal{A}_{\text{sa}})_1}^{\text{sot}} = (\mathcal{A}''_{\text{sa}})_1$;
- $\overline{(\mathcal{A})_1}^{\text{sot}} = (\mathcal{A}'')_1$;
- $\overline{(\mathcal{A}^+)_1}^{\text{sot}} = (\mathcal{A}''_+)_1$;
- *if $\mathcal{A}$ is unital,* $\overline{\mathcal{U}(\mathcal{A})}^{\text{sot}} \supset \mathcal{U}(\mathcal{A}'')$.

*Proof.* The first three inclusions $\subset$ are clear, since sot-limits preserve contractions, selfadjointness, and positivity. In light of Exercise 12.1.17, for unitaries we cannot expect in general the inclusion $\subset$. So we need to prove the four inclusions $\supset$.

Suppose that $T \in (\mathcal{A}''_{\text{sa}})_1$. By the Double Commutant Theorem there exists a net $\{T_j\} \subset \mathcal{A}$ with $T_j \xrightarrow{\text{sot}} T$. The function $f(t) = (t \wedge 1) \vee (-1)$, that is

$$f(t) = \begin{cases} 1, & t \geq 1 \\ t, & t \in (-1, 1) \\ -1, & t < -1 \end{cases}$$

is continuous and bounded, so it is strongly continuous by Lemma 12.3.16. This means, since $\sigma(T) \subset [-1, 1]$ and $\|f(T_k)\| \leq \|f\|_\infty = 1$ for all $k$, that $T = f(T) = \lim_k f(T_k) \in \overline{(\mathcal{A}_{\text{sa}})_1}^{\text{sot}}$.

For the positive part we can use the exact same argument but with the function $g(t) = (t \wedge 1) \vee 0$.

When $T \in (\mathcal{A}'')_1$, we consider

$$\tilde{T} = \begin{bmatrix} 0 & T \\ T^* & 0 \end{bmatrix} \in M_2(\mathcal{A})^{\text{sa}} \subset \mathcal{B}(\mathcal{H} \oplus \mathcal{H}).$$

From Corollary 12.3.11 we know that

$$M_2(\mathcal{A})'' = M_2(\mathcal{A}'').$$

As $\tilde{T}$ is selfadjoint and $\tilde{T} \in (M_2(\mathcal{A}''))_1$, by the first part of the proof there exists a net $\{\tilde{T}_\alpha\} \subset (M_2(\mathcal{A}))_1$ of selfadjoints with $\tilde{T}_\alpha \xrightarrow{\text{sot}} \tilde{T}$. Given $\xi \in \mathcal{H}$ we consider $\tilde{\xi} = \begin{bmatrix} 0 \\ \xi \end{bmatrix} \in \mathcal{H} \oplus \mathcal{H}$. We have that $\tilde{T}_\alpha \tilde{\xi} \to \tilde{T}\xi$. Looking at what happens in the first coordinate, we see that $(T_\alpha)_{12} \xrightarrow{\text{sot}} T$, so $T \in \overline{(\mathcal{A})_1}^{\text{sot}}$.

Finally, when $\mathcal{A}$ is unital and $U \in \mathcal{U}(\mathcal{A}'')$, by the Spectral Theorem (12.4.1; this is ahead in the text, but there is no coherence issue) there exists $T \in (\mathcal{A}''_{\text{sa}})_1$ with $U = e^{2\pi i T}$. By the first part of the proof, $T = \lim_{\text{sot}} T_j$ for a net $\{T_j\} \subset (\mathcal{A}_{\text{sa}})_1$. As $t \longmapsto e^{2\pi i t}$ is bounded and hence strongly continuous by Lemma 12.3.16, we get that $U = \lim_{\text{sot}} e^{2\pi i T_j} \in \overline{\mathcal{U}(\mathcal{A})}^{\text{sot}}$. $\square$

When $\mathcal{H}$ is separable both the sot and wot are metrizable on bounded sets (Exercise 12.1.3), and so the nets in Kaplansky's Density Theorem can be taken to be sequences.

Here is a much used consequence of the Double Commutant and Kaplansky's Density Theorems. We have not defined yet what it means for a representation to be irreducible as we don't need that notion yet, but it will be any of the three equivalent conditions below.

***12.3.18. Proposition.*** ***(Equivalent Definitions of Irreducibility)***

*Let $\mathcal{A}$ be a $C^*$-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. The following statements are equivalent:*

*(i)* $\pi(\mathcal{A})' = \mathbb{C}\, I_{\mathcal{H}}$*;*

*(ii)* $\pi(\mathcal{A})'' = \mathcal{B}(\mathcal{H})$*;*

*(iii) given* $\varepsilon > 0$*,* $T \in \mathcal{B}(\mathcal{H})$ *and* $\mathcal{K} \subset \mathcal{H}$ *a finite-dimensional subspace, there exists* $a \in \mathcal{A}$ *with* $\|a\| \leq \|T|_{\mathcal{K}}\|$ *and* $\|(\pi(a)-T)|_{\mathcal{K}}\| < \varepsilon$*. When $T$ is selfadjoint, the element $a$ may be chosen selfadjoint.*

*Proof.* (i) $\implies$ (ii) We have $\pi(\mathcal{A})'' = (\pi(\mathcal{A})')' = (\mathbb{C}\, I_{\mathcal{H}})' = \mathcal{B}(\mathcal{H})$.

(ii) $\implies$ (iii) The condition $\pi(\mathcal{A})'' = \mathcal{B}(\mathcal{H})$ guarantees that $\pi(\mathcal{A})$ is nondegenerate. By the Double Commutant Theorem, $\mathcal{B}(\mathcal{H}) = \pi(\mathcal{A})'' = \overline{\pi(\mathcal{A})}^{\mathrm{sot}}$. Let $\{\xi_1, \ldots, \xi_n\}$ be an orthonormal basis of $\mathcal{K}$, and put $\mathcal{W} = \{S \in \mathcal{B}(\mathcal{H}) : \|(T-S)\xi_j\| < \varepsilon/\sqrt{n},\ j = 1, \ldots, n\}$. This is a sot-neighbourhood of $T$, so by hypothesis there exists $a \in \mathcal{A}$ with $\pi(a) \in \mathcal{W}$. Given $\xi \in \mathcal{K}$ with $\|\xi\| = 1$, we can write $\xi = \sum_j c_k \xi_k$ with $\sum |c_k|^2 = 1$. Then

$$\begin{aligned}\|(\pi(a)-T)\xi\| &= \Big\|\sum_k c_k(\pi(a)-T)\xi_k\Big\| \leq \sum_k |c_k|\, \|(\pi(a)-T)\xi_k\| \\ &\leq \Big(\sum_k |c_k|^2\Big)^{1/2} \Big(\sum_k \|(\pi(a)-T)\xi_k\|^2\Big)^{1/2} < \varepsilon.\end{aligned}$$

If $T = T^*$, by Kaplansky's Density Theorem we can choose $\pi(a)$ selfadjoint, and if needed we replace $a$ with $\operatorname{Re} a$.

(iii) $\implies$ (i) The hypothesis gives us that $\overline{\pi(\mathcal{A})}^{\mathrm{sot}} = \mathcal{B}(\mathcal{H})$. By the Double Commutant Theorem,

$$\pi(\mathcal{A})' = (\pi(\mathcal{A})'')' = (\overline{\pi(\mathcal{A})}^{\mathrm{sot}})' = \mathcal{B}(\mathcal{H})' = \mathbb{C}\, I_{\mathcal{H}}. \qquad \square$$

In the proof of (ii) $\implies$ (iii) above it is tempting, when $T$ is selfadjoint, to simply replace $a$ with $\operatorname{Re} a$ without appealing to Kaplansky. The problem is that in general the operation of taking adjoint is not sot-continuous. Kaplansky allows us to take $\pi(a)$ to be selfadjoint, and then we use that $\pi(\operatorname{Re} a) = \operatorname{Re} \pi(a) = \pi(a)$.

We will use the following repeatedly.

***12.3.19. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra. Then $\overline{B_1^{\mathcal{M}}(0)}$ is wot-closed and $\mathcal{M}$ is $\sigma$-weak closed.*

*Proof.* By Corollary 12.1.3 we have $\overline{B_1^{\mathcal{M}}(0)}^{\text{wot}} = \overline{B_1^{\mathcal{M}}(0)}^{\text{sot}}$. Suppose that $T \in \overline{B_1^{\mathcal{M}}(0)}^{\text{sot}}$, $\{T_j\} \subset \mathcal{M}$, and $T_j \xrightarrow{\text{sot}} T$. By Kaplansky (Theorem 12.3.17) there exists a net $\{S_k\} \subset \overline{B_1^{\mathcal{M}}(0)}$ such that $S_k \xrightarrow{\text{sot}} T$. Hence $T \in \overline{B_1^{\mathcal{M}}(0)}$, since $\|T\xi\| = \lim_k \|S_k\xi\| \leq \|\xi\|$ for all $\xi \in \mathcal{H}$. This shows that

$$\overline{B_1^{\mathcal{M}}(0)}^{\text{wot}} = \overline{B_1^{\mathcal{M}}(0)}^{\text{sot}} = \overline{B_1^{\mathcal{M}}(0)}.$$

Knowing that the closed unit ball of $\mathcal{M}$ is $\sigma$-weak closed (Lemma 12.1.21), we get that all closed balls are $\sigma$-weak closed, and then Krein–Šmulian (Theorem 7.2.25) gives us that $\mathcal{M}$ is $\sigma$-weak closed. □

***12.3.20. Corollary.***

*The closed unit ball of a von Neumann algebra is wot and $\sigma$-weak-compact.*

*Proof.* We know from Proposition 12.3.19 that the closed unit ball of a von Neumann algebra $\mathcal{M}$ is wot-closed. Then it is wot-compact by Proposition 12.1.8 and therefore $\sigma$-weak-compact by Lemma 12.1.21. □

### 12.3.1. Exercises.

(12.3.1) Show that $\mathcal{B}(\mathcal{H})' = \mathcal{K}(\mathcal{H})' = \mathbb{C}\, I_{\mathcal{H}}$, and $(\mathbb{C}\, I_{\mathcal{H}})' = \mathcal{B}(\mathcal{H})$.

(12.3.2) Let $\mathcal{H} = \ell^2(\mathbb{N})$. Let $\mathcal{A} = \{M_a : a \in \ell^\infty(\mathbb{N})\}$ be the algebra of multipliers. Show that $\mathcal{A}' = \mathcal{A}$.

(12.3.3) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that

$$\mathcal{Z}(\mathcal{M}') = \mathcal{Z}(\mathcal{M}).$$

Conclude that $\mathcal{M}$ is a factor if and only if $\mathcal{M}'$ is a factor.

(12.3.4) Let $D \in M_n(\mathbb{C})$ be diagonal with all diagonal entries distinct. Show $\{D\}' = \{D\}'' = \{E \in M_n(\mathbb{C}) :\ \text{diagonal}\}$.

(12.3.5) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a finite-dimensional C$^*$-algebra. Show that $\mathcal{A}$ is a von Neumann algebra.

(12.3.6) Prove Proposition 12.3.2.

(12.3.7) Prove Proposition 12.3.4.

(12.3.8) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{P} \subset \mathcal{M}$ a set of projections. Show that $\bigvee \mathcal{P}$ and $\bigwedge \mathcal{P}$ are respectively the supremum and infimum of $\mathcal{P}$.

(12.3.9) Given an alternative proof of Corollary 12.3.10 by using an approximate unit.

(12.3.10) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate C*-algebra and $n \in \mathbb{N}$. Consider $M_n(\mathcal{A}) \subset \mathcal{B}(\mathcal{H}^n)$ and show that

$$M_n(\mathcal{A})' = \{X \otimes I_n : \ X \in \mathcal{A}'\}, \tag{12.4}$$

$$\{X \otimes I_n : \ X \in \mathcal{A}\}' = M_n(\mathcal{A}'), \tag{12.5}$$

$$M_n(\mathcal{A})'' = M_n(\mathcal{A}''). \tag{12.6}$$

(12.3.11) Let $X \subset \mathcal{B}(\mathcal{H})$. Prove that $W^*(X) = (X \cup X^*)''$, and show by example that it is possible to have $X'' \subsetneq W^*(X)$.

(12.3.12) Show that $\mathcal{S}$ and $\mathcal{S}^b$, as in Lemma 12.3.16, are closed under uniform limits.

(12.3.13) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\mathcal{M}_0 \subset \mathcal{M}$ a subspace. Show that the following statements are equivalent:

(i) $\mathcal{M}_0$ is wot-dense in $\mathcal{M}$;

(ii) $\mathcal{M}_0$ is sot-dense in $\mathcal{M}$;

(iii) $\mathcal{M}_0$ is $\sigma$-weak dense in $\mathcal{M}$.

## 12.4. The Spectral Theorem

We proved the Spectral Theorem for compact operators in Theorem 10.6.12. We are now ready to discuss the more general version of this, namely normal operators which are not necessarily compact. There is even a bit more generality available in that there is a Spectral Theorem for unbounded normal operators, but we will not discuss it here.

The version analog to Theorem 10.6.12 is Theorem 12.4.4 below. We will first prove the result that every normal operator is unitarily equivalent to a multiplication operator; so yes, it turns out that multiplication operators are pretty general!

The theorem deals with a general abelian C*-algebra $C_0(K)$ because the effort to prove the theorem for a single operator and an abelian C*-algebra that contains it, is the same.

***12.4.1. Theorem. (Spectral Theorem: Multiplication Operator Version)***

*Let $K$ be a locally compact Hausdorff space, and $\pi : C_0(K) \to \mathcal{B}(\mathcal{H})$ a unital $*$-homomorphism. Then there exists a measure space $X$, a unitary $W : \mathcal{H} \to L^2(X)$, and a unital $*$-homomorphism $\rho : C_0(K) \to L^\infty(X)$ such that*

$$\pi(f) = W^* M_{\rho(f)} W, \qquad f \in C_0(K).$$

*In particular, for any normal $T \in \mathcal{B}(\mathcal{H})$, there exists a measure space $X$, an $f \in L^\infty(X)$, and a unitary $W : \mathcal{H} \to L^2(X)$ with $T = W^* M_f W$.*

*Proof.* For each $\xi \in \mathcal{H}$ the map $f \longmapsto \langle \pi(f)\xi, \xi\rangle$ is a positive linear functional on $C_c(K)$. So Riesz–Markov (Theorem 2.9.2, we don't really need the more general version of Theorem 5.6.15 here) implies that there exists a unique regular Borel measure $\mu_\xi$ on $K$ such that

$$\langle \pi(f)\xi, \xi\rangle = \int_K f \, d\mu_\xi.$$

Define $V_\xi : C_c(K) \to \mathcal{H}$ by $V_\xi f = \pi(f)\xi$. We have

$$\|V_\xi f\|^2 = \|\pi(f)\xi\|^2 = \langle \pi(|f|^2)\xi, \xi\rangle = \int_X |f|^2 \, d\mu_\xi = \|f\|_2^2,$$

so $V$ is isometric if we consider $C_c(K) \subset L^2(K, \mu_\xi)$. As $C_c(K)$ is dense in $L^2(K, \mu_\xi)$ (Proposition 2.8.18), we can extend $V_\xi$ to an isometry of $L^2(K, \mu_\xi)$ into $\mathcal{H}$. If we restrict the codomain of $V_\xi$ to $\mathcal{H}_\xi = \overline{\pi(C_c(K))\xi}$, then $V_\xi$ is a unitary and $\xi$ is cyclic for the algebra $\pi(C_c(K))$.

We have, for any $f, g \in C_c(K)$,

$$V_\xi M_f \, g = V_\xi(fg) = \pi(fg)\xi = \pi(f)\pi(g)\xi = \pi(f)\, V_\xi g$$

Using again that $C_c(K)$ is dense in $L^2(K, \mu_\xi)$, we get that $V_\xi M_f = \pi(f) V_\xi$ for all $f \in C_c(K)$, and by continuity for all $f \in C_0(K)$. That is, $\pi(f) = V_\xi \, M_f \, V_\xi^*$. So we somehow got our goal, but our unitary does not have $\mathcal{H}$ as domain, but rather the subspace $\mathcal{H}_\xi$. To address this we do the following. If $\mathcal{H}_\xi \subsetneq \mathcal{H}$, we can take $\eta \in \mathcal{H}_\xi^\perp$ and repeat the above construction; note that $\mathcal{H}_\eta \perp \mathcal{H}_\xi$. Here we want to use Zorn's Lemma. Let $\mathcal{F}$ be the set of all families $\{\mathcal{H}_{\xi_j}\}_{j \in J}$ of pairwise orthogonal subspaces with $H_{\xi_j} = \overline{\pi(C_c(K))\xi_j}$. With the relation being inclusion, it is straightforward to see that any chain admits an upper bound: namely, the union of the families in the chain. Then Zorn's Lemma guarantees that $\mathcal{F}$ has a maximal element $\{\mathcal{H}_{\xi_j}\}_j$. Because this family is maximal, necessarily $[\bigoplus_j H_{\xi_j}]^\perp = \{0\}$; otherwise, we could repeat the construction above to get a new cyclic subspace and contradict the maximality. So $\mathcal{H} = \bigoplus_j H_{\xi_j}$ (this is a proper sum in $\mathcal{H}$, but we use the direct sum notation to emphasize the pairwise orthogonality of the subspaces). And then we can consider the unitary

$$W = \bigoplus_j V_{\xi_j}^* : \mathcal{H} \to \Big( \bigoplus_j L^2(K, \mu_{\xi_j}) \Big)_{\ell^2}.$$

Define $X = \bigcup_j K \times \{j\}$ (where we use the second entry to make it a disjoint union). Let

$$\Sigma = \{B \subset X : \ B \cap (K \times \{j\}) \text{ is a Borel subset of } K \times \{j\} \ \text{ for all } j\}.$$

This is a $\sigma$-algebra and

$$\mu(B) = \sum_j \mu_{\xi_j}\big(\pi_1(B \cap (K \times \{j\}))\big)$$

is a measure (Exercise 2.3.27). Then

$$\Big( \bigoplus_j L^2(K, \mu_{\xi_j}) \Big)_{\ell^2} \simeq L^2(X, \Sigma, \mu),$$

(Exercise 5.3.8) so we can think of $W : \mathcal{H} \to L^2(X, \mu)$. Letting $\rho(f) = (f)_j$,

$$\begin{aligned} W\pi(f) \sum_j \pi(g_j)\xi_j &= \sum_j V_{\xi_j}\pi(f)\pi(g_j))\xi_j = \sum_j M_f V_\xi \pi(g_j)\xi_j \\ &= M_{\rho(f)} W \sum_j \pi(g_j)\xi_j \end{aligned}$$

Thus $\pi(f) = W^* M_{\rho(f)} W$.

For the case where $T \in \mathcal{B}(\mathcal{H})$ is a normal operator, let $\pi = \Gamma^{-1} : C(\sigma(T)) \to \mathcal{B}(\mathcal{H})$ be the inverse of the Gelfand Transform (Theorem 11.2.3). We have shown that there exists a measure space $X$, $\rho : C(\sigma(T)) \to L^\infty(X)$, and a unitary $W : \mathcal{H} \to L^2(X)$ such that $\pi(f) = W^* M_{\rho(f)} W$. Taking $z$ to be the identity function $z(t) = t$, we have $\pi(z) = T$, so $T = W^* M_f W$, where $f = \rho(z)$. □

Here is a rather direct consequence of the Spectral Theorem that we promised on page 688.

### *12.4.2. Corollary.*

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Then $\overline{W(T)} = \overline{\operatorname{conv}}\sigma(T)$.*

*Proof.* We know from Proposition 10.2.3 that $\overline{W(T)} \supset \overline{\operatorname{conv}}\sigma(T)$. By Theorem 12.4.1 we may assume that $T = M_f$ acting on $L^2(X, \mu)$, where $f \in L^\infty(X, \mu)$. By Proposition 12.2.2 we know that $\sigma(M_f) = \operatorname{ess\,ran} f$. Let

$\lambda \in W(M_f)$. Then there exists $g \in L^2(X)$, with $\|g\|_2 = 1$, and

$$\lambda = \langle M_f g, g\rangle = \int_X f\,|g|^2\,d\mu.$$

Given $\varepsilon > 0$, we can form a simple function $s = \sum_k \lambda_k 1_{E_k}$ such that $\lambda_k \in \operatorname{ess\,ran} f$ for all $k$, and $\|f - s\|_\infty < \varepsilon$. Since the sets $\{E_k\}$ form a partition for $X$, the numbers

$$\alpha_k = \int_{E_k} |g|^2\,d\mu$$

are convex coefficients, i.e. $\alpha_k \geq 0$ for all $k$ and $\sum_k \alpha_k = 1$. Then

$$\Big|\lambda - \sum_k \alpha_k\lambda_k\Big| = \left|\int_X (f-s)\,|g|^2\,d\mu\right| \leq \|f - s\|_\infty < \varepsilon.$$

Thus $\lambda \in \overline{\operatorname{conv}}\operatorname{ess\,ran} f = \overline{\operatorname{conv}}\,\sigma(M_f) = \overline{\operatorname{conv}}\,\sigma(T)$. □

*12.4.3. Remark* . We can now give the straightforward proof of Proposition 10.3.7 promised on page 694. Indeed, if $T$ is normal with $\sigma(T) \subset \mathbb{R}$, by Corollary 12.4.2 we get $W(T) \subset \mathbb{R}$. Then Proposition 10.3.6 gives us that $T = T^*$.

We now turn to discuss spectral measures. Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Fix $\xi \in \mathcal{H}$ with $\|\xi\| = 1$. On the space $\mathbb{C}[x]$ of polynomials, consider the map $\varphi_\xi : f \longmapsto \langle f(T)\xi, \xi\rangle$. We can see the polynomial $f$ as an element of $C(\sigma(T))$; what this means for us, in practice, is that it establishes the norm of $f$ as $\|f\| = \sup\{|f(t)| :\ t \in \sigma(T)\}$. By Proposition 9.2.9 and Corollary 12.4.2, we have that $\varphi_\xi(f) \in \overline{\operatorname{conv}} f(\sigma(T))$. It follows that

$$|\varphi_\xi(f)| \leq \max\{|f(t)| :\ t \in \sigma(T)\} = \|f\|.$$

So $\varphi_\xi$ is bounded and by density it can be uniquely extended to all of $C(\sigma(T))$ (no need for Hahn–Banach here: Exercise 5.5.11; alternatively, we could have used the continuous functional calculus from the start). By Riesz–Markov there exists a unique regular complex Borel measure $\mu_\xi$ such that

$$\varphi_\xi(f) = \int_{\sigma(T)} f(\lambda)\,d\mu_\xi(\lambda), \quad f \in C(\sigma(T)). \tag{12.7}$$

Since $\varphi_\xi(f) \in \overline{\operatorname{conv}} f(\sigma(T))$ for all $f$, we get from Exercise 2.10.18 that $\mu_\xi$ is a positive measure. Now fix a Borel subset $E \subset \sigma(T)$, and define a form on $\mathcal{H}$ by

$$[\xi, \eta]_E = \frac{1}{4}\sum_{k=0}^{3} i^k\,\mu_{\xi + i^k\eta}(E). \tag{12.8}$$

The uniqueness also gives us that $\mu_{\alpha\xi} = |\alpha|^2\,\mu_\xi$, and using this property and evaluating we see that $[\xi, \xi]_E = \mu_\xi(E)$. The form satisfies the Paralellogram Identity;

indeed, recalling that $\sigma(T)$ is compact, and combining Proposition 2.8.18 with Theorem 7.4.20 and Corollary 2.8.13, we can get a sequence of polynomials $\{f_j\}$ that are uniformly bounded on $E$, say $\|f_j\|_\infty < c$, and such that $f_j \to 1_E$ pointwise. Then, using Dominated Convergence and (12.7) repeatedly,

$$\begin{aligned}[]
[\xi+\eta,\xi+\eta]_E + [\xi-\eta,\xi-\eta]_E &= \mu_{\xi+\eta}(E) + \mu_{\xi-\eta}(E)\\
&= \lim_j \varphi_{\xi+\eta}(f_j) + \varphi_{\xi-\eta}(f_j)\\
&= \lim_j \langle f_j(T)(\xi+\eta),\xi+\eta\rangle + \langle f_j(T)(\xi-\eta),\xi-\eta\rangle\\
&= \lim_j 2\langle f_j(T)\xi,\xi\rangle + 2\langle f_j(T)\eta,\eta\rangle\\
&= \lim_j 2\varphi_\xi(f_j) + 2\varphi_\eta(f_j)\\
&= 2\mu_\xi(E) + 2\mu_\eta(E) = 2[\xi,\xi] + 2[\eta,\eta].
\end{aligned}$$

Using Dominated Convergence again

$$\begin{aligned}
|\mu_\xi(E) - \mu_\eta(E)| &= \left| \int_{\sigma(T)} 1_E\, d\mu_\xi - \int_{\sigma(T)} 1_E\, d\mu_\eta \right|\\
&= \lim_j \left| \int_{\sigma(T)} f_j\, d\mu_\xi - \int_{\sigma(T)} f_j\, d\mu_\eta \right|\\
&= \lim_j |\varphi_\xi(f_j) - \varphi_\eta(f_j)| = \lim_j |\langle f_j(T)\xi,\xi\rangle - \langle f_j(T)\eta,\eta|\\
&= \lim_j |\langle f_j(T)\xi,\xi-\eta\rangle + \langle f_j(T)(\xi-\eta),\eta\rangle|\\
&\leq c\,(\|\xi\| + \|\eta\|)\,\|\xi-\eta\|,
\end{aligned}$$

showing the continuity of $\xi \longmapsto \mu_\xi(E)$. Now Remark 4.2.9 applies to give us that $[\cdot,\cdot]$ is a positive sesquilinear form. Knowing that

$$[\xi,\xi]_E = \mu_\xi(E) \leq \mu_\xi(\sigma(T)) = \varphi_\xi(1) = \|\xi\|^2,$$

Cauchy–Schwarz implies that $[\cdot,\cdot]$ is bounded. Then Proposition 10.1.5 gives us a unique operator $\mu(E) \in \mathcal{B}(\mathcal{H})$ such that $[\xi,\eta] = \langle \mu(E)\xi,\eta\rangle$ for all $\xi,\eta \in \mathcal{H}$. Since $\langle \mu(E)\xi,\xi\rangle = [\xi,\xi]_E \geq 0$ for all $\xi \in \mathcal{H}$, $\mu(E)$ is positive; in particular it is selfadjoint (Proposition 10.3.6). As we can repeat the process for any Borel set $E \subset \sigma(T)$, we have a map $\mu$, defined on the Borel subsets of $T$, and taking values in $\mathcal{B}(\mathcal{H})_+$.

If $\{f_r\}$ is a sequence of polynomials as above (that is, uniformly bounded) with $f_r \to 1_E$ pointwise, then Dominated Convergence gives

$$\langle \mu(E)\xi,\xi\rangle = \int_{\sigma(T)} 1_E\, d\mu_\xi = \lim_r \int_{\sigma(T)} f_r\, d\mu_\xi = \lim_r \varphi_\xi(f_r) = \lim_r \langle f_r(T)\xi,\xi\rangle. \tag{12.9}$$

So the sequence $\{f_r(T)\}$, via polarization, converges wot to $\mu(E)$. Let $E,F \subset \sigma(T)$ be Borel sets and $\{f_j\},\{g_j\} \subset C[0,1]$ be bounded sequences of polynomials with $f_j \to 1_E$ and $g_j \to 1_F$ pointwise. Then, using Dominated Convergence once more,

$$\begin{aligned}\langle \mu(E\cap F)\xi,\xi\rangle &= \mu_\xi(E\cap F) = \int_{\sigma(T)} 1_{E\cap F}\,d\mu_\xi = \int_{\sigma(T)} 1_E\,1_F\,d\mu_\xi \\ &= \lim_k \lim_j \int_{\sigma(T)} f_j\,g_k\,d\mu_\xi = \lim_k\lim_j \langle f_j g_k(T)\xi,\xi\rangle \\ &= \lim_k\lim_j \langle f_j(T)g_k(T)\xi,\xi\rangle = \lim_k \langle \mu(E)g_k(T)\xi,\xi\rangle \\ &= \lim_k \langle g_k(T)\xi,\mu(E)\xi\rangle = \langle \mu(F)\xi,\mu(E)\xi\rangle = \langle \mu(E)\mu(F)\xi,\xi\rangle.\end{aligned}$$

As this can be done for any $\xi\in\mathcal{H}$, via polarization we obtain that

$$\mu(E\cap F) = \mu(E)\mu(F). \tag{12.10}$$

In particular $\mu(E) = \mu(E\cap E) = \mu(E)^2$, so $\mu(E)$ is a projection for all Borel $E\subset\sigma(T)$. We say that $\mu$ is a **spectral measure** for $T$. To check that it behaves as a measure, if $\{E_k\}\subset\sigma(T)$ are Borel and pairwise disjoint,

$$\begin{aligned}\Big\langle \mu\Big(\bigcup_k E_k\Big)\xi,\xi\Big\rangle &= \mu_\xi\Big(\bigcup_k E_k\Big) = \sum_k \mu_\xi(E_k) = \sum_k \langle \mu(E_k)\xi,\xi\rangle \\ &= \Big\langle \sum_k \mu(E_k)\xi,\xi\Big\rangle.\end{aligned} \tag{12.11}$$

The last step—which is less trivial than it looks, for $\sum_k \mu(E_k)$ is an operator, a sot limit of sums of pairwise orthogonal projections—is justified by Proposition 12.1.10, since $\sum_{k=1}^m \mu(E_k) = \mu(\bigcup_{k=1}^m E_k) \leq \mu(\sigma(T))$. This last argument requires the additivity of $\mu$, which can be first obtained by doing (12.11) over finitely many sets.

For any simple function $\sum_k \lambda_k\,1_{E_k}$, with $\lambda_k\in E_k$ and $\{E_1,\ldots,E_m\}$ a Borel partition of $\sigma(T)$,

$$\int_{\sigma(T)} \sum_k \lambda_k\,1_{E_k}\,d\mu_\xi = \sum_k \lambda_k\,\mu_\xi(E_k) = \sum_k \lambda_k\langle \mu(E_k)\xi,\xi\rangle = \Big\langle \sum_k \lambda_k\mu(E_k)\xi,\xi\Big\rangle$$

If we choose the simple functions $f_r = \sum_k \lambda_{k,r}\,1_{E_{k,r}}$ so that they are uniformly bounded and they converge monotonically to the identity function $\mathrm{id}:\lambda\longmapsto\lambda$, we get

$$\langle T\xi,\xi\rangle = \int_{\sigma(T)} \lambda\,d\mu_\xi(\lambda) = \lim_r \int_{\sigma(T)} f_r\,d\mu_\xi = \lim_r \Big\langle \sum_k \lambda_{k,r}\mu(E_{k,r})\xi,\xi\Big\rangle. \tag{12.12}$$

The limit on the right allows us to represent $T$ as an operator valued integral

$$T = \int_{\sigma(T)} \lambda\,d\mu. \tag{12.13}$$

At first sight, it might look like a stretch to interpret (12.12) as (12.13). But since the convergence $f_r(\lambda) \to \lambda$ can be made uniform, it follows that

$$\left|\left\langle (T - \sum_k \lambda_{k,r}\mu(E_{k,r}))\xi, \xi\right\rangle\right| = \left|\left\langle \sum_k \int_{E_k} (\lambda - \lambda_{k,r})\mu(E_{k,r})\xi, \xi\right\rangle\right|$$
$$\leq \|\mathrm{id} - f_r\| \, \langle \mu(\sigma(T))\xi, \xi\rangle = \|\mathrm{id} - f_r\| \, \|\xi\|^2,$$

so the convergence in (12.13) is actually in norm and not just wot. That is,

$$\int_{\sigma(T)} \lambda \, d\mu = \lim_r \sum_k \lambda_{k,r}\, \mu(E_{k,r}),$$

as a norm limit.

In summary, we have proven

***12.4.4. Theorem. (Spectral Theorem: Spectral Measure Version)***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Then there exists a unique spectral measure $\mu_T$ on the Borel subsets of $\sigma(T)$ such that*

$$T = \int_{\sigma(T)} \lambda \, d\mu_T(\lambda).$$

*Proof.* All that remains to be shown is that $\mu$ is unique. We will defer said proof until after Corollary 12.4.11. □

*12.4.5. Remark*. It is fairly common in the literature to denote the spectral measure of $T$ by $E$, or $E_T$, $E^T$, etc. We prefer not to do so to avoid confusion with the common use of $E$ for measurable subsets. The only drawback of using $\mu_T$ for the spectral measure is the possibility of confusion between a scalar-valued and an operator-valued integral; the distinction should not be hard within the context.

*12.4.6. Remark* . Theorem 10.6.12 is a particular case of Theorem 12.4.4 (Exercise 12.4.1).

*12.4.7. Example.* Let $\mathcal{H} = L^2[0,1]$ with Lebesgue measure and $T$ the multiplication operator $(Tf)(t) = tf(t)$. Then $\sigma(T) = [0,1]$. Let us calculate $\mu_T$. From the arguments after (12.8), we have

$$\langle \mu_T(E)\xi, \xi\rangle = \lim_r \langle f_r(T)\xi, \xi\rangle$$

for a sequence of polynomials $f_r$ with $f_r \to 1_E$ pointwise. Being polynomials, it is immediate that $(f_r(T)g)(t) = f_r(t)g(t)$. So $f_r(T)g \to 1_E\, g$ pointwise. Being bounded and in a context of finite measure, Dominated Convergence gives us that $f_r(T) \xrightarrow{\text{wot}} M_{1_E}$ (the multiplication operator by $1_E$). Then $\mu_T(E) = M_{1_E}$. See Exercise 12.4.3 for the general case of this.

*12.4.8. Remark* . While not said explicitly neither in the statement of the Spectral Theorem nor in the proof, we always have that $\mu_T(\sigma(T)) = I_{\mathcal{H}}$ (Exercise 12.4.7).

**12.4.1. Borel Functional Calculus.** Given $K \subset \mathbb{C}$ compact, we consider

$$B_b(K) = \{f : K \to \mathbb{C} : \text{ Borel, such that } \sup\{|f(t)| : t \in K\} < \infty\},$$

the algebra of bounded Borel functions on $K$. The operations are pointwise addition and multiplication, pointwise multiplication by scalars, and conjugation as the adjoint.

**12.4.9. Definition.**

*Given two Borel measures $\mu, \nu$ on $\mathbb{C}$, we write $\mu \bowtie \nu$ if $\mu \ll \nu$ and $\nu \ll \mu$.*

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal and let $\mu_T$ be its spectral measure. A* ***scalar spectral measure*** *is a measure $\mu$ on the Borel subsets of $\sigma(T)$ such that $\mu \bowtie \mu_T$.*

An easy way to obtain a scalar spectral measure for $T$ is to fix an orthonormal basis $\{\xi_n\}$ and put

$$\mu(E) = \sum_n \alpha_n \langle \mu_T(E)\xi_n, \xi_n\rangle = \sum_n \alpha_n \|\mu_T(E)\xi_n\|^2$$

for some positive weights with $\alpha \in \ell^1(\mathbb{N})$. This choice guarantees that $\mu$ is finite. The only non-trivial part of checking that $\mu$ is a measure is the $\sigma$-additivity, but that follows easily by Tonelli and the $\sigma$-additivity of $\mu_T$ as in (12.11). Of course this construction only works if $\mathcal{H}$ is separable, which justifies the restriction in Corollary 12.4.11 below. Better said, the construction of $\mu$ can be always made, but $\mu$ may not be finite, or it could even fail to be semifinite.

**12.4.10. Definition. (Borel Functional Calculus)**

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal and $f : \sigma(T) \to \mathbb{C}$ a bounded Borel function. We define $f(A) \in \mathcal{B}(\mathcal{H})$ to be the operator*

$$f(A) = \int_{\sigma(T)} f(\lambda)\, d\mu_T(\lambda).$$

While the separability and $\sigma$-finiteness restrictions in Corollary 12.4.11 below may seem a bit forced (at least if a scalar spectral measure $\mu$ exists), if $\mu$ is too wild then $L^\infty(\mu)$ may fail to be a von Neumann algebra; concretely, $\mu$ needs to be localizable (see Remark 5.6.12 and recall that in a von Neumann algebra families of projections admit a supremum by Corollary 12.3.9).

***12.4.11. Corollary. (Properties of Borel Functional Calculus)***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $T \in \mathcal{M}$ be normal, with $\mathcal{H}$ separable. Fix a $\sigma$-finite scalar spectral measure $\mu$ for $T$. Then $W^*(T) \simeq L^\infty(\mu)$ via a (wot, weak$^*$)-continuous $*$-isomorphism that maps $T$ to the identity function $\mathrm{id} : z \longmapsto z$. The map $f \longmapsto f(T)$ is a $*$-homomorphism regardless of the existence of $\mu$ or the separability of $\mathcal{H}$.*

*Proof.* By Exercise 2.10.14 and the fact that $L^\infty$ is defined in terms of nullsets, we may assume without loss of generality that $\mu$ is finite.

The process to define the integral leading to (12.13) works not only for the identity function, but for any bounded Borel function. So we can define $\Gamma : B_b(\sigma(T)) \to W^*(T)$ by

$$\Gamma(f) = \int_{\sigma(T)} f(\lambda)\, d\mu_T(\lambda). \tag{12.14}$$

This map is linear (because integrals are) and $*$-preserving (because the measure is projection-valued, hence positive). The Spectral Theorem gives us that $\Gamma(\mathrm{id}) = T$. So the only nontrivial part in showing that $\Gamma$ is a $*$-homomorphism is multiplicativity. Suppose that $f, g$ are simple. Then

$$f = \sum_k \lambda_k\, 1_{E_k}, \qquad g = \sum_j \gamma_j 1_{F_j}$$

for finite measurable partitions $\{E_k\}$ and $\{F_j\}$ of $\sigma(T)$. Using (12.10),

$$\begin{aligned}\int_{\sigma(T)} f(\lambda)g(\lambda)\, d\mu_T(\lambda) &= \int_{\sigma(T)} \sum_{k,j} \lambda_k\gamma_j 1_{E_k}(\lambda)1_{F_j}(\lambda)\, d\mu_T(\lambda)\\ &= \sum_{k,j} \lambda_k\gamma_j \int_{\sigma(T)} 1_{E_k\cap F_j}(\lambda)\, d\mu_T(\lambda)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k,j} \lambda_k \gamma_j \, \mu_T(E_k \cap F_j) = \sum_{k,j} \lambda_k \gamma_j \, \mu_T(E_k) \, \mu_T(F_j) \\
&= \Big( \sum_k \lambda_k \mu_T(E_k) \Big) \Big( \sum_j \gamma_j \, \mu_T(F_j) \Big) \\
&= \Big( \int_{\sigma(T)} f(\lambda) \, d\mu_T(\lambda) \Big) \Big( \int_{\sigma(T)} g(\lambda) \, d\mu_T(\lambda) \Big).
\end{aligned}
$$

Given arbitrary bounded Borel $f, g$, we may take sequences of simple functions that converge uniformly to each (Theorem 2.4.13), and then apply Dominated Convergence the way it was done in the build up to (12.13), to get that

$$
\begin{aligned}
\Gamma(fg) &= \int_{\sigma(T)} f \, g \, d\mu_T = \Big( \int_{\sigma(T)} f \, d\mu_T \Big) \Big( \int_{\sigma(T)} g \, d\mu_T \Big) \\
&= \Gamma(f)\Gamma(g), \qquad f, g \in B_b(\sigma(T)).
\end{aligned}
$$

If $f = g$ $\mu$-a.e., then $f(T) = g(T)$ by definition of integral. So it makes sense to consider the domain of $\Gamma$ to be $L^\infty(\mu)$.

If $f : \sigma(T) \to \mathbb{C}$ satisfies that $\mu(\{|f| > 0\}) \neq 0$, there has to exist $\delta > 0$ such that $\mu(\{|f| > \delta\}) \neq 0$ (from $\{|f| > 0\} = \bigcup_n \{|f| > \frac{1}{n}\}$). Then $\mu_T(\{|f| > \delta\} \neq 0$ and

$$
|f|(T) \geq \int_{\{|f| > \delta\}} |f(\lambda)| \, d\mu_T(\lambda) \geq \delta \mu_T(\{|f| > \delta\}),
$$

so $|f|(T) \neq 0$. This means that if $\Gamma(f) = 0$, then $|f|(T) = \Gamma(|f|) = |\Gamma(f)| = 0$ and therefore $|f| = 0$ $\mu$-a.e. Thus $\Gamma$ is injective.

Next we need to argue that the range of $\Gamma$ falls inside $W^*(T)$. The argument in (12.9) shows that $\mu_T(E) = \lim_{\text{sot}} f_j(T)$ for appropriately chosen polynomials $f_j$. This means that $\mu_T(E) \in \overline{W^*(T)}^{\text{sot}} = W^*(T)$. As taking linear combinations and norm-limits will stay within $W^*(T)$, we have that $\Gamma(f) \in W^*(T)$ for all $f \in L^\infty(\mu)$.

Let us now look at the continuity. Suppose that $f_j \xrightarrow{\text{weak}^*} 0$ in $L^\infty(\mu)$. From $\mu_\xi(E) = \langle \mu_T(E)\xi, \xi \rangle$ and $\mu \bowtie \mu_T$ we see that $\mu_\xi \ll \mu$. Then by Theorem 2.10.10 there exists a Radon–Nikodym derivative $h \in L^1(\mu)$. For $\xi \in \mathcal{H}$,

$$
\langle \Gamma(f_j)\xi, \xi \rangle = \int_{\sigma(T)} f_j \, d\mu_\xi = \int_{\sigma(T)} f_j \, h \, d\mu \to 0.
$$

After using polarization, $\Gamma(f_j) \xrightarrow{\text{wot}} 0$. Conversely, suppose that $\Gamma(f_j) \xrightarrow{\text{wot}} 0$. Assume momentarily that $W^*(T)$ has a separating vector $\xi$; the name "separating" is still to be defined, but we mean that $S\xi = 0$ implies $S = 0$ for all $S \in W^*(T)$. Let $V : L^2(\mu_\xi) \to \mathcal{H}$ be initially given by $Vf = \Gamma(f)\xi$ for

$f \in L^\infty(\mu_\xi)$. Since

$$\|\Gamma(f)\xi\|^2 = \langle \Gamma(|f|)\xi, \xi\rangle = \int_{\sigma(T)} |f|^2 \, d\mu_\xi = \|f\|^2,$$

we have that $V$ is an isometry. As $L^\infty(\mu_\xi)$ is dense in $L^2(\mu_\xi)$ (Proposition 2.8.16), $V$ extends to an isometry. As the range of $V$ contains $p(T,T^*)\xi$ for every $p \in \mathbb{C}[x,y]$ and those vectors are dense in $\mathcal{H}$ (because $\{p(T,T^*) : p\}$ is sot-dense in $W^*(T)$), $V$ is unitary. Fix $h \in L^1(\mu)$, positive. Then $h\mu \ll \mu_\xi$; indeed, if $\mu_\xi(E) = 0$, this is $\langle \mu_T(E)\xi, \xi\rangle = 0$, and then $\mu_T(E) = 0$ by the hypothesis on $\xi$. Hence $\mu(E) = 0$ and $h\mu(E) = 0$. Let $g$ be the Radon–Nikodym derivative $g = d(h\mu)/d\mu_\xi \in L^1(\mu_\xi)$ (this is where we need that $\mu$ be finite). So $g^{1/2} \in L^2(\mu_\xi)$. Define $\nu = Vg^{1/2}$. Now

$$\begin{aligned} \int_{\sigma(T)} f_j h \, d\mu &= \int_{\sigma(T)} f_j g \, d\mu_\xi = \langle f_j g^{1/2}, g^{1/2}\rangle \\ &= \langle \Gamma(f_j)\, V g^{1/2}, V g^{1/2}\rangle = \langle \Gamma(f_j)\nu, \nu\rangle \to 0 \end{aligned}$$

(note that $g^{1/2} = \lim g_n$ with $g_n$ bounded, and $V(f_j g_n) = \Gamma(f_j)\Gamma(g_n)\xi = \Gamma(f_j)Vg_n$ and then take limit). Therefore $f_j \xrightarrow{\text{weak}^*} 0$. To finish the continuity argument we need to see that $\xi$ exists. We use Zorn's Lemma to construct orthonormal $\{\xi_n\} \subset \mathcal{H}$ with $\mathcal{H} = \bigoplus_n W^*(T)\xi_n$ and the subspaces $\{W^*(T)\xi_n\}$ pairwise orthogonal (countably many since $\mathcal{H}$ is separable, this is the place in the proof where we need the hypothesis). Then $\xi = \sum_n \frac{1}{n}\xi_n$ satisfies the requirement; indeed, noting that the projections $P_n$ onto $\overline{W^*(T)\xi_n}$ are in $W^*(T)'$ by Corollary 12.3.14, if $S \in W^*(T)$ with $S\xi = 0$, then for all $R \in W^*(T)$ we have $SR\xi_n = nSRP_n\xi = nRPS\xi = 0$. So $S = 0$ on each $\overline{W^*(T)\xi_n}$ and therefore $S = 0$. This finishes the argument that $\Gamma$ is bicountinuous.

It remains to check that $\Gamma$ is surjective. The algebra $W^*(T)$ is the smallest von Neumann algebra that contains $T$. As already mentioned, $W^*(T)$ is the sot-closure of $\{p(T,T^*) : p \in \mathbb{C}[x,y]\}$. As $p(T,T^*) = \Gamma(p)$ (from $\Gamma(z) = T$ and $\Gamma$ a $*$-homomorphism), we have that for any $S \in W^*(T)$ there exist polynomials $\{p_j\} \subset \mathbb{C}[x,y]$ with $p_j(T,T^*) \xrightarrow{\text{wot}} S$. By the previous paragraph, $\{p_j\}$ is weak$^*$-Cauchy in $L^\infty(\mu)$ so it converges to some $f$; and then the continuity of $\Gamma$ gives us $\Gamma(f) = \lim_{\text{sot}} \Gamma(p_j) = \lim_{\text{sot}} p_j(T,T^*) = S$. Hence $\Gamma$ is surjective. □

As it was mentioned in previous versions of functional calculus and used in the proof of Corollary 12.4.11, the fact that the map $\Gamma$ is a $*$-homomorphism and the fact that it maps the identity to $T$, immediately give us $\Gamma(p) = p(T)$ for any polynomial, and because $\Gamma$ is norm-continuous (Proposition 11.4.9) we also get $\Gamma(f) = f(T)$ for all $f \in C(\sigma(T))$. So the Borel functional calculus does extend the continuous functional calculus. For non-continuous Borel $f$, the new functional calculus gives us $f(T)$, and that becomes a definition of $f(T)$;

a reasonable definition since it agrees with previous notions where the contexts overlap, it maps pointwise operations between functions to the corresponding operations in $W^*(T)$, and it behaves nicely with respect to the natural topologies. In particular, for any Borel subset $E \subset \sigma(T)$,

$$1_E(T) = \mu_T(E). \tag{12.15}$$

Indeed, if we approximate $1_E$ with a uniformly bounded sequence of continuous functions $f_r$ as in the process leading to (12.13), then for $\xi \in \mathcal{H}$

$$\langle 1_E(T)\xi, \xi\rangle = \int_{\sigma(T)} 1_E \, d\mu_\xi = \mu_\xi(E) = \langle \mu_T(E)\xi, \xi\rangle.$$

As this can be done for any $\xi \in \mathcal{H}$, we obtain (12.15).

*12.4.12. Remark*. The separability requirement in Corollary 12.4.11 is not a technicality. Let $\mathcal{H} = \ell^2[0,1]$ and $T = M_z$ the multiplication operator by the identity function. Let $f : [0,1] \to \mathbb{C}$ be any bounded function, say $|f| \leq c$. Let $\mathcal{F} = \{F \subset [0,1] : \text{ finite}\}$, ordered by inclusion. For each $F \in \mathcal{F}$ choose a polynomial $p_F$ such that $p_F(t) = f(t)$ for all $t \in F$, and $\|p_F\|_\infty \leq 2c$. Fix $\varepsilon > 0$ and $\xi \in \mathcal{H}$. Then $\xi = \sum_{t\in R} c_t \, \delta_t$, where $R \subset [0,1]$ is countable. Choose $R_0 \subset R$, finite, such that $\sum_{t \in R\setminus R_0} |c_t|^2 < \varepsilon^2$. For any $F \in \mathcal{F}$ with $F \supset R_0$, and with $\xi_0 = \sum_{t\in R_0} c_t\delta_t$,

$$\begin{aligned}
\|p_F(M_z)\xi - M_f\xi\| = \|M_{p_F - f}\xi\| &\leq \|M_{p_F-f}(\xi - \xi_0)\| + \|M_{p_F-f}\xi_0\| \\
&\leq (\|M_{p_F}\| + \|M_f\|)\varepsilon + \Big(\sum_{t\in R_0} |p_F(t) - f(t)|^2|c_t|^2\Big)^{1/2} \\
&\leq 3c\varepsilon.
\end{aligned}$$

This shows that $M_{p_F} \xrightarrow{\text{sot}} M_f$, which implies that $W^*(M_z) = \{M_f : f : [0,1] \to \mathbb{C} \text{ bounded}\}$. This is strictly larger than $L^\infty[0,1]$. Indeed, fix a non-measurable set $E$ and consider $f = 1_E$. Then, for any $g \in L^\infty[0,1]$, there exists $t \in [0,1]$ and $\delta > 0$ with $|1_E(t) - g(t)| \geq \delta$. Then $\|(M_{1_E} - M_g)\delta_t\| \geq \delta$, which means that $\|M_{1_E} - M_g\| \geq \delta$. So $M_f$ does not belong to $L^\infty[0,1]$ when the latter is seen as multiplication operators on $L^2[0,1]$.

*Proof of uniqueness in Theorem 12.4.4.* Suppose we have a projection-valued measure $\mu'$ on the Borel subsets of $\sigma(T)$ such that

$$T = \int_{\sigma(T)} \lambda \, d\mu'(\lambda).$$

The proof of Corollary 12.4.11 shows that the map $f \longmapsto \int_{\sigma(T)} f(\lambda)\,d\mu_T$ is linear and multiplicative, and the same happens for $\mu'$. Therefore

$$\int_{\sigma(T)} f\,d\mu' = \int_{\sigma(T)} f\,d\mu_T \tag{12.16}$$

for polynomials $f$. Given $E \subset \sigma(T)$ Borel, we choose polynomials $f_r$ as in page 859, uniformly bounded on $\sigma(T)$ and with $f_r \to 1_E$ pointwise. Then $f_r \to 1_E$ weak* in $L^\infty$ (both for a scalar-valued spectral measure for $\mu_T$ and a scalar-valued spectral measure for $\mu'$) by dominated convergence. Then Corollary 12.4.11 implies that

$$\begin{aligned}\mu'(E) &= \int_{\sigma(T)} 1_E\,d\mu' = \lim_r \int_{\sigma(T)} f_r\,d\mu' \\ &= \lim_r \int_{\sigma(T)} f_r\,d\mu = \int_{\sigma(T)} 1_E\,d\mu_T = \mu_T(E).\end{aligned}$$

Thus $\mu' = \mu_T$. □

Note that we cannot use (12.15) directly to prove the uniqueness of $\mu_T$, because a priori we do not know that $1_E(T)$ is the same operator whether we define it with $\mu_T$ or with $\mu'$.

**12.4.13. Corollary. (Spectral Mapping, Borel Version)**

*Let $\mathcal{M}$ be a von Neumann algebra, $T \in \mathcal{M}$ be normal, and $f : \sigma(T) \to \mathbb{C}$ Borel. Then*

$$\sigma(f(T)) = \operatorname{ess\,ran} f. \tag{12.17}$$

*Proof.* By the isomorphism in Corollary 12.4.11, $\sigma(f(T)) = \sigma(M_f)$. Then Proposition 12.2.2 finishes the proof. □

As opposed to the previous versions of spectral mapping, we cannot include—though it would look prettier—$f(\sigma(T))$ in (12.17). The reason is that $f$ may take values on nullsets that become irrelevant for the functional calculus.

**12.4.14. Corollary.**

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\lambda_0 \in \mathbb{C}$. Consider the extension of $\mu_T$ to all of $\mathbb{C}$ as in Exercise 2.3.8. The following statements are equivalent:*

*(i)* $\lambda_0 \in \sigma(T)$;

*(ii)* $\mu_T(B_r(\lambda_0)) \neq 0$ *for all* $r > 0$.

*Proof.* (ii) $\implies$ (i) Suppose that $\lambda_0 \notin \sigma(T)$. As $\sigma(T)$ is closed, there exists $r > 0$ with $B_r(\lambda_0) \subset \mathbb{C} \setminus \sigma(T)$. Then $\mu_T(B_r(\lambda_0)) = 0$ by the construction of the extension of $\mu_T$ as in Exercise 2.3.8.

(i) $\implies$ (ii) Suppose that there exists $r > 0$ with $\mu_T(B_r(\lambda_0)) = 0$. Then $\mu_T(\sigma(T) \setminus B_r(\lambda_0)) = I_{\mathcal{H}}$. That is, $T = T\,1_{\sigma(T)\setminus B_r(\lambda_0)}(T)$. Since $|\lambda - \lambda_0| \geq r$ for all $\lambda \in \sigma(T) \setminus B_r(\lambda_0)$, we can define $S = f(T)$, where $f$ is the bounded Borel function $f(\lambda) = \frac{1}{\lambda - \lambda_0}\,1_{\sigma(T)\setminus B_r(\lambda_0)}(\lambda)$. Then, as

$$T - \lambda_0 I_{\mathcal{H}} = T\,1_{\sigma(T)\setminus B_r(\lambda_0)}(T) - \lambda_0\,1_{\sigma(T)\setminus B_r(\lambda_0)}(T) = g(T),$$

with $g(\lambda) = (\lambda - \lambda_0)\,1_{\sigma(T)\setminus B_r(\lambda_0)}(\lambda)$,

$$(T - \lambda_0 I)S = 1_{\sigma(T)\setminus B_r(\lambda_0)}(T) = I_{\mathcal{H}}.$$

As $S$ and $T$ commute, we have that $\lambda_0 \notin \sigma(T)$. □

***12.4.15. Corollary.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, and $T \in \mathcal{M}$ be normal. Let $f : \sigma(T) \to \mathbb{C}$ be a bounded Borel function. Then $f(T) \in \mathcal{M}$. In particular, the projections onto the kernel and closure of the range of $T$ are in $\mathcal{M}$.*

*Proof.* By the construction in Corollary 12.4.11, $f(T) \in W^*(T) \subset \mathcal{M}$. The projections onto $\ker T$ and $\overline{\operatorname{ran}} T$ are of the same form, by Exercise 12.4.8. □

The next two corollaries are used very regularly in von Neumann algebra technicalities.

***12.4.16. Corollary.***

*Let $\mathcal{M}$ be a von Neumann algebra. Then $\mathcal{M} = \overline{\operatorname{span}}^{\|\cdot\|}\,\mathcal{P}(\mathcal{M})$.*

*Proof.* Looking at real and imaginary parts we have $\mathcal{M} = \operatorname{span}\mathcal{M}^{\rm sa}$, so it is enough to show that a selfadjoint $T \in \mathcal{M}^{\rm sa}$ is a norm-limit of linear combinations of projections. The Spectral Theorem implies that $T$ is a norm limit of linear combinations of projections of the form $\mu_T(E)$. And by Corollary 12.4.15, $\mu_T(E) \in \mathcal{M}$ for all Borel $E$. □

***12.4.17. Corollary.***

*Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}^+$ nonzero. Then there exist $\lambda > 0$ and a projection $P \in \mathcal{P}(\mathcal{M})$ with $PT = TP$ and $\lambda P \leq PT$.*

*Proof.* Exercise 12.4.6. □

***12.4.18. Corollary.***

*Let $\mathcal{H}$ be a separable Hilbert space and $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ an abelian von Neumann algebra. Then $\mathcal{A}$ is singly generated; that is, there exists $T \in \mathcal{A}^+$ such that $\mathcal{A} = W^*(T)$.*

*Proof.* As explained in page 903, the hypothesis implies that the algebra $\mathcal{A}$ is sot-separable. So there exist countably many $\{T_n\} \subset \mathcal{A}$ with $\mathcal{A} = \overline{\{T_n\}}^{\text{sot}}$. By writing $T_n = \operatorname{Re} T_n + i\operatorname{Im} T_n$ and relabelling, we may assume that $T_n = T_n^*$ for all $n$ and $\mathcal{A} = \overline{\{T_n\}}^{\text{sot}}$. By the Spectral Theorem for each $n, k \in \mathbb{N}$ there exist projections $P_{n,k,1}, \ldots, P_{n,k,r_{n,k}}$ such that

$$\operatorname{dist}(T_n, \operatorname{span}\{P_{n,k,1}, \ldots, P_{n,k,r_{n,k}}\}) < \frac{1}{k}.$$

This implies that $\mathcal{A} = W^*(\{P_{n,k,s}\})$. After relabelling again, $\mathcal{A} = W^*(\{P_n\})$. Now let

$$T = \sum_{n=1}^\infty 3^{-n} P_n.$$

By construction, $T \in \overline{\operatorname{span}}^{\|\cdot\|} \{P_n : n\} \subset \mathcal{A}$. Using that $\mathcal{A}$ is abelian,

$$0 \leq (I_\mathcal{A} - P_1)T = (I_\mathcal{A} - P_1)\sum_{n=2}^\infty 3^{-n} P_n \leq (I_\mathcal{A} - P_1)\sum_{n=2}^\infty 3^{-n} I_\mathcal{A} = \frac{1}{6}(I_\mathcal{A} - P_1).$$

We also have

$$P_1 T \geq \frac{1}{3} P_1, \qquad P_1 T = \sum_{n=1}^\infty 3^{-n} P_1 P_n \leq \sum_{n=1}^\infty 3^{-n} P_1 = \frac{1}{2} P_1.$$

This means that $T = P_1 T + (I_\mathcal{A} - P_1)T$, pairwise orthogonal operators. By Exercise 9.5.7, $\sigma(T) \subset \left[0, \frac{1}{6}\right] \cup \left[\frac{1}{3}, \frac{1}{2}\right]$. Because the two components of $T$ have product zero among them, we have $f(T) = f(P_1 T) + f((I_\mathcal{A} - P_1)T)$ for all $f \in C(\sigma(T))$ (we could use Borel functional calculus here, but we don't really need it). If we take $f = 1_{[1/3,1/2]}$ (note that $f \in C(\sigma(T))$), $f(T) = f(P_1 T)$. Let $Q = f(T)$. This is a projection with $\frac{1}{3} Q \leq QT \leq \frac{1}{2} Q$ and $0 \leq (I_\mathcal{A} - Q)T \leq \frac{1}{6}(I_\mathcal{A} - Q)$ (note that $1 - f = 1_{[0,1/6]}$ on $C(\sigma(T))$). By Exercise 10.5.15, $Q = P_1$. Thus $P_1 \in C^*(T)$. Now we have that $\sum_{n=2}^\infty 3^{-n} P_n = T - \frac{1}{3} P_1 \in C^*(T)$, and we can rehash the argument to get $P_2 \in C^*(T)$. Inductively, $P_n \in C^*(T)$ for all $n$, and therefore $\mathcal{A} \subset W^*(T)$. But $T \in \mathcal{A}$, so $\mathcal{A} = W^*(T)$. □

***12.4.19. Corollary.***

*Let $T \in \mathcal{B}(\mathcal{H})$ be normal, and let $\varepsilon > 0$. Then there exists a Borel function $f$ on $\sigma(T)$ such that $f(T)$ is invertible and $\|T - f(T)\| < \varepsilon$. In particular, any normal operator is a norm-limit of normal invertible operators.*

*Proof.* By Corollary 12.4.11, it is enough to show that there exists $f \in L^\infty(\mu)$, invertible, with $\|\mathrm{id}_{\sigma(A)} - f\|_\infty < \varepsilon$. Using Theorem 2.4.13, there exists $g$ simple with $\|\mathrm{id}_{\sigma(A)} - g\|_\infty < \varepsilon/2$. Since $g$ is finite-valued, there exists $\delta > 0$ with $\delta < \varepsilon$ and $f = g + \delta$ invertible. Then

$$\|\mathrm{id}_{\sigma(A)} - f\|_\infty \leq \|\mathrm{id}_{\sigma(A)} - g\|_\infty + \|g - f\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

When talking about matrices, the Spectral Theorem makes it obvious that if $R$ is normal and $RT = TR$, then $R^*T = TR^*$ (because the spectral projections can be written as polynomials on $R$; or even easier, $R^*$ itself can be written as a polynomial in $R$). The general case cannot be obtained directly from the Spectral Theorem, though. The problem is that when $R$ is not selfadjoint the polynomials one uses in the Spectral Theorem to approximate the spectral measure include both $R$ and $R^*$, so we cannot conclude that $T$ commutes with $\mu_R(E)$. The result we need, that $RT = TR$ implies $R^*T = TR^*$ if $R$ is normal, was proven by Fuglede [Fug50]. The result is commonly stated in the generalized form proven by Putnam [Put51].

***12.4.20. Theorem. (Fuglede–Putnam)***

*Let $R, S \in \mathcal{B}(\mathcal{H})$ be normal, $T \in \mathcal{B}(\mathcal{H})$. If $RT = TS$, then $R^*T = TS^*$.*

*Proof.* Assume first that $R = S$. Let $g : \mathbb{C} \to \mathcal{B}(\mathcal{H})$ be $g(z) = e^{izR^*}Te^{-izR^*}$. Since $e^{izR^*} = \sum_{k=0}^\infty \frac{(iR^*)^k}{k!} z^k$, it is clear that it is differentiable. It is thus (operator-valued) analytic. It follows by the usual arguments that $g$ is differentiable. If the reader is uncomfortable with the idea of doing operator-valued complex analysis, it is always possible to resort to the trick we used in the proof of Theorem 9.2.12, to compose $g$ with a bounded linear functional and get a regular analytic function. Whichever avenue one prefers, the usual rules of differentiation apply and one deduces that $g'(z) = iR^*g(z) - g(z)(iR^*)$. In particular

$$g'(0) = i(R^*T - TR^*).$$

From $RT = TR$ we get that $e^{izR}T = Te^{izR}$. Then

$$g(z) = e^{izR^*}Te^{-izR^*} = e^{izR^*}e^{i\bar{z}R}Te^{-i\bar{z}R}e^{-izR^*} = e^{2i\mathrm{Re}\,zR}Te^{-2i\mathrm{Re}\,zR},.$$

where the normality of $R$ is what allows us to do $e^{izR^*}e^{i\bar{z}R} = e^{izR^*+i\bar{z}R}$ (Exercise 9.3.3). The fact that $\operatorname{Re} zR$ is selfadjoint makes $e^{2i\operatorname{Re} zR}$ a unitary; then $\|g(z)\| = \|T\|$ for all $z$. As $g$ is bounded and entire, $g(z)$ is constant by Liouville's Theorem (Corollary 3.7.4; one either proves the operator-valued version, or uses it for $\varphi \circ g$ for each $\varphi \in \mathcal{B}(\mathcal{H})^*$). Thus $g'(0) = 0$, which means that $R^*T = TR^*$.

The argument above can be tweaked to make it work with $R$ and $S$. But we can also apply it to the operators

$$\tilde{R} = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix}, \qquad \tilde{T} = \begin{bmatrix} 0 & T \\ 0 & 0 \end{bmatrix}$$

on $\mathcal{B}(\mathcal{H} \oplus \mathcal{H})$, where $\tilde{R}$ is normal, to show that $RT = TS$ implies $R^*T = TS^*$. □

Historically, Putnam's extension of Fuglede's result allowed him to prove the following.

**12.4.21. Corollary.**

*Let $S, T \in \mathcal{B}(\mathcal{H})$ be normal. If $T$ and $S$ are similar, then they are unitarily equivalent.*

*Proof.* By hypothesis there exists $X \in \mathcal{B}(\mathcal{H})$, invertible, with $XS = TX$. By Fuglede–Putnam, $XS^* = T^*X$. Then

$$X^*XS = X^*TX = SX^*X.$$

Iterating this we obtain that $p(X^*X)S = Sp(X^*X)$ for all polynomials $p$, and by taking norm limits the equality still holds for any continuous function $p \in C(\sigma(S))$. In particular,

$$|X|S = S|X|.$$

Let $W|X| = X$ be the polar decomposition of $X$ (Proposition 10.4.11). Since $X$ is invertible, $W$ is a unitary. Now

$$WS|X| = W|X|S = XS = TX = TW|X|.$$

Since $|X|$ is invertible, we get $WS = TW$, which is $WSW^* = T$. □

**12.4.2. Exercises.**

(12.4.1) Let $T \in \mathcal{K}(\mathcal{H})$ be normal. Show that Theorem 10.6.12 is a particular case of Theorem 12.4.4.

(12.4.2) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that there exists $S \in \mathcal{B}(\mathcal{H})$, selfadjoint, and a continuous $f : \sigma(S) \to \mathbb{C}$ such that $T = f(S)$. *(Hint: Proposition 7.6.7)*

(12.4.3) Expanding on the ideas of Example 12.4.7, show that if $g \in L^\infty[0,1]$, then $\mu_{M_g}(E) = M_{1_{g^{-1}(E)}}$.

(12.4.4) Show that the extreme points in the convex set $\mathcal{B}(\mathcal{H})_1^+$ of positive operators with norm at most 1, are precisely the projections.

(12.4.5) Let $T \in \mathcal{B}(\mathcal{H})$ be normal, $\lambda_0 \in \mathbb{C}$. Consider the extension of $\mu_T$ to all of $\mathbb{C}$ as in Exercise 2.3.8. Show that the following statements are equivalent:

 (i) $T$ is compact;

 (ii) for every $\lambda_0 \in \sigma(T) \setminus \{0\}$ there exists $r > 0$ such that $\mu_T(B_r(\lambda_0))$ is a finite-rank projection.

(12.4.6) Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}^+$ nonzero. Show that there exists a nonzero projection $P \in \mathcal{M}$ that commutes with $T$ and $\lambda > 0$ such that $TP \geq \lambda P$.

(12.4.7) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that the construction of $\mu_T$ in the proof of the Spectral Theorem gives $\mu_T(\sigma(T)) = I_{\mathcal{H}}$.

(12.4.8) Let $T \in \mathcal{B}(\mathcal{H})$ be normal. Show that $1_{\{0\}}(T)$ is the projection onto $\ker T$, and that $1_{\sigma(T)\setminus\{0\}}(T)$ is the projection onto $\overline{\operatorname{ran}}\, T$.

(12.4.9) Let $\mathcal{M} = \ell^\infty[0,1] \subset \mathcal{B}(L^2[0,1])$, seen as multiplication operators. Fix $t \in [0,1]$ and for $\delta > 0$ let $P_\delta = M_{1_{[t-\delta,t+\delta]}}$.

 (i) Show that $P_\delta \xrightarrow{\text{sot}} 0$ as $\delta \to 0$.

 (ii) Show that if $K \in \mathcal{K}(L^2[0,1])$ then $\|P_\delta K\| \to 0$.

 (iii) Show that there exists $T \in \mathcal{B}(L^2[0,1])$ such that $P_\delta T$ does not converge in norm to 0.

(12.4.10) Show that in Corollary 12.4.18, if a sot-dense separable C*-subalgebra $\mathcal{A}_0 \subset \mathcal{A}$ is prescribed, the operator $T$ can be chosen so that $C^*(T) \supset \mathcal{A}_0$. *(Attention: the word "separable" has different meanings when referring to C* and von Neumann algebras, see subsection 12.7.2)*

(12.4.11) Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}$ normal. Give an alternative proof of Corollary 12.3.8 by using Corollary 12.4.15 and the fact that $UTU^* = T$ for all unitaries $U \in \mathcal{M}'$.

## 12.5. Cyclic and Separating Vectors

We have already defined cyclic vectors in Definition 11.6.1, but the emphasis there was in the representation, while here it is on the algebra.

**12.5.1. Definition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. A vector $\xi \in \mathcal{H}$ is **cyclic** for $\mathcal{M}$ if $\overline{\mathcal{M}\xi} = \mathcal{H}$. And $\xi$ is **separating** for $\mathcal{M}$ if for each $X \in \mathcal{M}$, $X\xi = 0$ implies $X = 0$. Cyclic vectors are often called **generating**.*

***12.5.2. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $\xi \in \mathcal{H}$. Then*

*(i) $\xi$ is separating for $\mathcal{M}$ if and only if it is cyclic for $\mathcal{M}'$.*

*(ii) $\xi$ is separating for $\mathcal{M}'$ if and only if it is cyclic for $\mathcal{M}$.*

*Proof.* Suppose that $\xi$ is cyclic for $\mathcal{M}'$. If $T \in \mathcal{M}$ and $T\xi = 0$, then $TS\xi = ST\xi = 0$ for all $S \in \mathcal{M}'$. That is, $T\mathcal{M}'\xi = 0$. As $\mathcal{M}'\xi$ is dense in $\mathcal{H}$, we get that $T = 0$, and so $\xi$ is separating for $\mathcal{M}$. Conversely suppose that $\xi$ is not cyclic for $\mathcal{M}'$. Then there exists nonzero $\eta \in (\mathcal{M}'\xi)^\perp$. Let $P$ be the orthogonal projection onto $\overline{\mathcal{M}'\xi}$. By Corollary 12.3.14, $P \in \mathcal{M}$. Then $I_\mathcal{M} - P \in \mathcal{M}$ is nonzero (since $(I_\mathcal{M} - P)\eta = \eta$) and $(I_\mathcal{M} - P)\xi = 0$, so $\xi$ is not separating for $\mathcal{M}$. This shows (i).

Now (ii) is, via the Double Commutant Theorem, (i) for the von Neumann algebra $\mathcal{M}'$. ☐

***12.5.3. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a $*$-subalgebra. Then there exists a pairwise orthogonal family $\{\xi_j\} \subset \mathcal{H}$ such that $\mathcal{M}\xi_j \perp \mathcal{M}\xi_k$ for $k \neq j$, and*

$$\mathcal{H} = \left( \bigoplus_j \overline{\mathcal{M}\xi_j} \right)_{\ell^2}.$$

*Proof.* We will use Zorn's Lemma. Let

$$\mathcal{F} = \{(\eta_j)_{j\in J_\alpha} : \ \text{ with } \mathcal{M}\eta_j \perp \mathcal{M}\eta_k \ \text{ for } k \neq j\},$$

the set of all families as in the statement, whether they cover $\mathcal{H}$ or not. $\mathcal{F} \neq \varnothing$ because $(\mathcal{M}\eta) \in \mathcal{F}$ for any $\eta \in \mathcal{H}$. We order $\mathcal{F}$ by inclusion. If $\{(\eta_j)_{j\in J_\alpha}\}_\alpha$ is a chain, then by taking $J = \bigcup_\alpha J_\alpha$ we get $\{\eta_j\}_{j\in J}$ as an upper bound. By Zorn's Lemma there exists a maximal $\{\xi_j\}_{j\in J}$. Suppose that

$$\Big(\bigoplus_j \overline{\mathcal{M}\xi_j}\Big)_{\ell^2} \subsetneq \mathcal{H}.$$

Let $\xi \in \left(\bigoplus_j \overline{\mathcal{M}\xi_j}\right)^\perp$. For any $j \in J$, we have $\langle \xi, T\xi_j\rangle = 0$ for all $T \in \mathcal{M}$. Then, for any $S \in \mathcal{M}$,

$$\langle S\xi, T\xi_j\rangle = \langle \xi, S^*T\xi_j\rangle = 0.$$

Thus $\mathcal{M}\xi \perp \mathcal{M}\xi_j$ for all $j$, contradicting the maximality. □

**12.5.4. Definition.**

*A projection $P \in \mathcal{M}$ is **σ-finite** if any pairwise orthogonal family of subprojections of $P$ is countable. The von Neumann algebra $\mathcal{M}$ is **σ-finite** if the identity $I_\mathcal{M}$ is σ-finite.*

When $\mathcal{H}$ is separable, every $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is $\sigma$-finite. And if $\mathcal{M} = \mathcal{B}(\mathcal{H})$, $\sigma$-finiteness is equivalent to $\mathcal{H}$ separable. Some authors use the term **countably decomposable** instead of $\sigma$-finite.

**12.5.5. Definition.**

*Let $(X, \mu)$ be a measure space. We define*

$$\mathcal{A}_\mu = \{M_f : \ f \in L^\infty(\mu)\} \subset \mathcal{B}(L^2(X,\mu)).$$

*A **masa** (maximal abelian subalgebra) in a von Neumann algebra $\mathcal{M}$ is an abelian $*$-subalgebra $\mathcal{A} \subset \mathcal{M}$ such that $\mathcal{A} \subset \mathcal{B} \subset \mathcal{M}$ with $\mathcal{B}$ abelian implies $\mathcal{B} = \mathcal{A}$.*

If $\mathcal{A}$ is a masa in $\mathcal{B}(\mathcal{H})$, then $\mathcal{A}' \subset \mathcal{A}$ since otherwise we would be able to enlarge $\mathcal{A}$ while keeping it abelian. As we also have $\mathcal{A} \subset \mathcal{A}'$ from $\mathcal{A}$ abelian, $\mathcal{A} = \mathcal{A}'$ and a fortiori $\mathcal{A} = \mathcal{A}''$. This last equality can also be seen from the fact that the wot-closure of an abelian $*$-algebra is an abelian von Neumann algebra. Bottom line is, a $*$-algebra $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is a masa if and only if $\mathcal{A} = \mathcal{A}'$, and this forces also $\mathcal{A} = \mathcal{A}''$.

The next result has way more general versions, but we will only state and prove what we need.

***12.5.6. Proposition.***

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a $\sigma$-finite abelian von Neumann algebra. Then $\mathcal{A}$ has a separating vector $\xi$. If $\mathcal{A}$ is maximal abelian, then $\xi$ is also cyclic.*

*Proof.* By Proposition 12.5.3 and $\sigma$-finiteness there exists a countable family $\{\xi_n\} \subset \mathcal{H}$ such that $\mathcal{A}'\xi_n \perp \mathcal{A}'\xi_m$ for all $n \neq m$, and $\mathcal{H} = \bigoplus_n \overline{\mathcal{A}'\xi_n}$. We may assume, without loss of generality, that $\|\xi_n\| = 1$ for all $n$. Let $\xi = \sum_n \frac{1}{n}\,\xi_n$. This converges in $\mathcal{H}$ because the $\xi_n$ are orthonormal. If $P_n$ is the orthogonal projection onto $\overline{\mathcal{A}'\xi_n}$, we have $P_n \in \mathcal{A}$ by Corollary 12.3.14. Also, $\sum_n P_n = I$. We want to show that $\xi$ is separating for $\mathcal{A}$. If $A \in \mathcal{A}$ and $A\xi = 0$, for any $B \in \mathcal{A}'$ we have, using that $\xi_n = nP_n\xi$,

$$AP_n(B\xi_n) = BAP_n\xi_n = nBP_nA\xi = 0.$$

Thus $AP_n\,\overline{\mathcal{A}'\xi_n} = 0$, giving us $AP_n = 0$. As this can be done for all $n$, we get that $A = \sum_n AP_n = 0$, and $\xi$ is separating.

When $\mathcal{A}$ is maximal abelian, by Proposition 12.5.2 we have that $\xi$ is cyclic for $\mathcal{A}' = \mathcal{A}$. $\square$

*12.5.7. Remark.* The $\sigma$-finiteness requirement is not just a technical restriction, but an intrinsic one. In the sense that if a masa $\mathcal{A} \subset \mathcal{H}$ has a cyclic/separating vector, then it is $\sigma$-finite (Exercise 12.5.4).

***12.5.8. Proposition.***

*Let $(X, \mu)$ be a $\sigma$-finite measure space. Then*

$$\mathcal{A}_\mu = \mathcal{A}_\mu' = \mathcal{A}_\mu''.$$

*That is, $\mathcal{A}_\mu$ is a masa in $\mathcal{B}(L^2(\mu))$. If $X = \mathbb{C}$, $\mu$ is finite and* $\operatorname{supp}\mu$ *is compact, then $\{M_z^\mu\}' = \mathcal{A}_\mu$.*

*Proof.* We always have, since $\mathcal{A}_\mu$ is abelian, $\mathcal{A}_\mu \subset \mathcal{A}_\mu'$. So it is enough to show that $\mathcal{A}_\mu' \subset \mathcal{A}_\mu$.

Suppose first that $\mu(X) < \infty$. Let $T \in \mathcal{A}_\mu'$. Put $g = T1$ ($1 \in L^2(\mu)$ since $\mu$ is finite). For any $f \in L^2(\mu)$,

$$Tf = TM_f1 = M_fT1 = M_fg = M_gf.$$

If $E_n = \{|g| \geq n\}$, then

$$\mu(E_n) = \frac{1}{n^2}\int_{E_n} n^2\,d\mu \leq \frac{1}{n^2}\int_{E_n} |g|^2\,d\mu = \frac{1}{n^2}\int_X |g\,1_{E_n}|^2\,d\mu$$

$$= \frac{1}{n^2} \int_X |T1_{E_n}|^2 \, d\mu = \frac{1}{n^2} \|T1_{E_n}\|^2$$
$$\leq \frac{\|T\|^2 \, \|1_{E_n}\|^2}{n^2} = \frac{\|T\|^2 \, \mu(E_n)}{n^2}.$$

This means that if $\mu(E_n) > 0$, then $\|T\| \geq n$. So $\mu(E_n) = 0$ for all $n > \|T\|$. Thus $g$ is essentially bounded, which means $g \in L^\infty(\mu)$ and $T = M_g \in \mathcal{A}_\mu$; this shows that $\mathcal{A}_\mu' \subset \mathcal{A}_\mu$.

Now suppose that $\mu(X) = \infty$ and let $T \in \mathcal{A}_\mu'$. For any $E \subset X$, measurable with $\mu(E) < \infty$, for any $f$ supported on $E$ we have $Tf = T1_E f = 1_E Tf \in L^2(\mu|E)$. That is, $T$ acts on $L^2(\mu|_E)$. By the previous paragraph there exists $g_E \in L^2(\mu|_E)$ such that $T = M_{g_E}$ on $L^2(\mu|_E)$. If $F \subset X$ is another measurable set with $E \subset F$ and $\mu(F) < \infty$, then

$$g_F \, 1_E = 1_E \, g_F \, 1_E = 1_E \, T1_E = g_E \, 1_E.$$

Since $X = \bigcup_n X_n$ with $\mu(X_n) < \infty$ for all $n$ and $X_n \subset X_{n+1}$, we can define $g$ by $g|_{X_n} = g_{X_n}$. Then $g$ is well-defined and measurable and, using Proposition 12.2.2,

$$\big\|g|_{X_n}\big\|_\infty = \big\|g_{X_n}\big\|_\infty = \big\|M_{g_{X_n}}\big\| = \|T|_{L^2(\mu|_{X_n})}\| \leq \|T\|.$$

So $\|g\|_\infty \leq \|T\|$, giving us $g \in L^\infty(\mu)$ and $T = M_g$. We have shown that $\mathcal{A}_\mu' = \mathcal{A}_\mu$, and it follows automatically that $\mathcal{A}_\mu' = \mathcal{A}_\mu''$.

Finally, consider the case where $X = \mathbb{C}$ and $\mu$ is finite and compactly supported. From $M_z^\mu \in \mathcal{A}_\mu$ we have $\mathcal{A}_\mu = \mathcal{A}_\mu' \subset \{M_z^\mu\}'$. Conversely, if $T \in \{M_z^\mu\}'$, by Fuglede–Putnam (Theorem 12.4.20) we have that $T$ also commutes with $M_{\bar z}^\mu$. It follows that $TM_f^\mu = M_f^\mu T$ for all polynomials $f(z, \bar z)$. These polynomials are norm-dense in the continuous functions, since we can work inside a compact set that contains $\operatorname{supp} \mu$, by Stone–Weierstrass (Theorem 7.4.20). As the compactly supported continuous functions are dense in $L^2(\mu)$ (Proposition 2.8.18, and note that regularity is granted by Proposition 2.9.11), we have $f = \lim_n p_n$ in $L^2(\mu)$ for polynomials $p_n$. Then

$$TM_f^\mu \, g = T(fg) = \lim_n T(p_n g) = \lim_n TM_{p_n}^\mu \, g = \lim_n M_{p_n}^\mu Tg$$
$$= \lim_n p_n Tg = f \, Tg = M_g^\mu T \, g.$$

Thus $T \in \mathcal{A}_\mu'$ and $\{M_z^\mu\}' = \mathcal{A}_\mu' = \mathcal{A}_\mu$. $\square$

In the particular case where $\mathcal{H} = L^2[0,1]$ with Lebesgue measure, we now know that $L^\infty[0,1]$, seen as multiplication operators, is a masa $\mathcal{A}_m$ in $\mathcal{B}(\mathcal{H})$. Here is another way to construct a masa. Fix an orthonormal basis $\{\xi_n\}$ for a separable Hilbert space $\mathcal{H}$, and let $\mathcal{A}_a$ (the a is for "atomic") be the diagonal operators:

$$\mathcal{A}_a = \{T \in \mathcal{B}(\mathcal{H}) : T\xi_n \in \mathbb{C}\xi_n \ \text{ for all } n\}. \tag{12.18}$$

Then (Exercise 12.5.3)

- $\mathcal{A}_a$ is a masa;
- $\mathcal{A}_a \simeq \ell^\infty(\mathbb{N})$;
- there exists a $*$-monomorphism $\mathcal{A}_a \to \mathcal{A}_m$;
- there exists no $*$-isomorphism between $\mathcal{A}_a$ and $\mathcal{A}_m$.

There are no other kinds of masas of $\mathcal{B}(\mathcal{H})$ (see Proposition 12.5.10 below for the concrete statement). Any masa in $\mathcal{B}(L^2[0,1])$ is isomorphic to one of $\mathcal{A}_a$ and $\mathcal{A}_m$. Masas of certain of von Neumann algebras, on the other hand, can be difficult to understand.

*12.5.9. Remark*. Similarly to the situation in Remark 12.5.7, the requirement that $\mu$ is $\sigma$-finite in Proposition 12.5.8 is not just a technicality. Consider the measure space $(X, \Sigma, \mu)$ where $X$ is an uncountable set, $\Sigma = \{A \subset X : A \text{ or } X \setminus A \text{ is countable}\}$ and $\mu(A) = \infty$ unless $A = \varnothing$. This is not a very interesting measure, but still $L^\infty(X)$ is big. Fix $X_0 \subset X$ uncountable and with uncountable complement, and let $B = \{1_A \subset X_0 : A \text{ is countable}\} \subset L^\infty(X)$. Then $B$ has no least upper bound in $L^\infty(X)$, as this would be $1_{X_0}$ which is not measurable. This shows that $L^\infty(X)$ is not a von Neumann algebra, since a monotone increasing net of selfadjoints converges sot and von Neumann algebras are sot-closed. The algebra $L^\infty(X)$ is a von Neumann algebra precisely when $\mu$ is localizable (see Remark 5.6.12).

Tying together results we have,

***12.5.10. Proposition.***

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a $\sigma$-finite abelian von Neumann algebra. The following statements are equivalent:*

*(i) there exists a regular Borel measure on $[0,1]$ such that $\mathcal{A}$ is unitarily equivalent to $\mathcal{A}_\mu$;*

*(ii) $\mathcal{A}$ is a masa;*

*(iii) $\mathcal{A}'$ is abelian;*

*(iv) $\mathcal{A}$ has a cyclic vector.*

*Proof.* (i) $\Longrightarrow$ (ii) We know that $\mathcal{A}_\mu$ is a masa from Proposition 12.5.8. The unitary isomorphism extends to an isomorphisms between $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(L^2(\mu))$, so $\mathcal{A}$ is a masa.

(ii) $\Longrightarrow$ (iii) As has been mentioned, since $\mathcal{A}$ is maximal abelian, $\mathcal{A}' = \mathcal{A}$.

(iii) $\Longrightarrow$ (iv) By hypothesis $\mathcal{A} \subset \mathcal{A}' \subset \mathcal{A}'' = \mathcal{A}$ is $\sigma$-finite. Hence it has a separating vector by Proposition 12.5.6. And then $\mathcal{A}$ has a cyclic vector by Proposition 12.5.2.

(iv) $\Longrightarrow$ (i) Let $\xi \in \mathcal{H}$ be cyclic for $\mathcal{A}$. By Corollary 12.4.18 there exists $T \in \mathcal{A}^+$ with $\mathcal{A} = W^*(T)$. By Theorem 11.2.3 there is an isomorphism $\pi : C(\sigma(T)) \to C^*(T) \subset \mathcal{A}$ with $\pi(z) = T$. Applying the first part of the proof of Theorem 12.4.1 to $\pi$ and $\sigma(T)$, there exists a regular Borel measure $\mu$ on $\sigma(T)$ and a unitary $V : \mathcal{H} \to L^2(\mu)$ such that $\pi(f) = VM_fV^*$. So we have $T = VM_z^\mu V^*$. As the map $S \longmapsto VSV^*$ is a normal $*$-isomorphism onto its range,

$$\begin{aligned} W^*(T) &= \overline{\{p(T) : \ p \in \mathbb{C}[x]\}}^{\text{sot}} = \overline{\{p(VM_z^\mu V^*) : \ p \in \mathbb{C}[x]\}}^{\text{sot}} \\ &= V\overline{\{p(M_z^\mu) : \ p \in \mathbb{C}[x]\}}^{\text{sot}} V^* = V\mathcal{A}_\mu V^*. \end{aligned}$$ □

Next is a result we will need later.

***12.5.11. Lemma.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $\xi \in \mathcal{H}$ a cyclic vector for $\mathcal{M}$, and $\eta \in \mathcal{H}$. Then there exist $\zeta \in \overline{S\mathcal{H}}$ and $S, T \in \mathcal{M}$, with $S \geq 0$, such that $S\zeta = \xi$ and $T\zeta = \eta$.*

*Proof.* By scaling if needed, we may assume that $\|\xi\| \leq 1$ and $\|\eta\| \leq 1$. Because $\xi$ is cyclic for $\mathcal{M}$ there exists a sequence $\{S_k'\} \subset \mathcal{M}$ with $S_k'\xi \to \eta$. By passing to a subsequence if needed, we may assume that $\|(S_k' - S_{k-1}')\xi\| < 4^{-k}$ for all $k$. Let $S_0' = 0$ and $S_k = S_k' - S_{k-1}' \in \mathcal{M}$. Then

$$\sum_{k=1}^{\infty} S_k\xi = \eta$$

and

$$\Big\|\eta - \sum_{k=1}^{n} S_k\xi\Big\| \leq \sum_{k=n+1}^{\infty} \|S_k\xi\| \leq \sum_{k=n+1}^{\infty} 4^{-k} = \frac{4^{-n-1}}{1-4^{-1}} = \frac{1}{3 \times 4^n} < \frac{1}{4^n}.$$

We define

$$H_n = \Big(I_{\mathcal{M}} + \sum_{k=1}^{n} 4^k S_k^* S_k\Big)^{1/2}.$$

This is an increasing sequence of positive invertible elements in $\mathcal{M}$ with $H_n \geq I_{\mathcal{M}}$ for all $n$. Hence $\{H_n^{-1}\}$ is a decreasing sequence of positive operators in $\mathcal{M}$ with $H_n^{-1} \leq I_{\mathcal{M}}$ for all $n$. Applying Proposition 12.1.10 to the sequence $\{I_{\mathcal{M}} - H_n^{-1}\}$, there exists $S_0 = \lim_{\text{sot}} I_{\mathcal{M}} - H_n^{-1} \in \mathcal{M}$. Thus $\{H_n^{-1}\}$ converges sot to $S = I_{\mathcal{M}} - S_0 \in \mathcal{M}$.

We have

$$\|H_n\xi\|^2 = \langle H_n^2\xi, \xi\rangle = \|\xi\|^2 + \sum_{k=1}^{n} 4^k \langle S_k^* S_k\xi, \xi\rangle$$

$$
\begin{aligned}
&= \|\xi\|^2 + \sum_{k=1}^{n} 4^k \|S_k \xi\|^2 \\
&\le 1 + \sum_{k=1}^{n} 4^k \left( \Big\| \eta - \sum_{j=1}^{k} S_j \xi \Big\| + \Big\| \sum_{j=1}^{k-1} S_j \xi - \eta \Big\| \right)^2 \\
&\le 1 + \sum_{k=1}^{n} 4^k \left( \frac{1}{4^k} + \frac{1}{4^{k-1}} \right)^2 \le 1 + \sum_{k=1}^{n} \frac{4^k\, 4}{4^{2k-2}} \\
&\le 1 + \sum_{k=1}^{n} \frac{1}{4^{k-3}} = 1 + 64\, \frac{4^{-1} - 4^{-n-1}}{1 - 4^{-1}} \le 1 + \frac{64}{3}.
\end{aligned}
$$

Therefore the sequence $\{H_n \xi\}$ is bounded in $\mathcal{H}$. By Banach–Alaoglu there exists a weak cluster point $\zeta \in \mathcal{H}$.

Given $\nu \in \mathcal{H}$ and $\varepsilon > 0$, choose $n$ such that

$$
\big|\langle \zeta - H_n \xi, S\nu \rangle\big| < \frac{\varepsilon}{2}, \qquad \text{and} \qquad \big\|(S - H_n^{-1})\nu\big\| < \frac{\varepsilon}{2 \sup\{\|H_m \xi\| :\ m\}}.
$$

Then

$$
\begin{aligned}
\big|\langle (S\zeta - \xi, \nu \rangle\big| &= \big|\langle \zeta - H_n \xi + H_n \xi, S\nu \rangle - \langle \xi, \nu \rangle\big| \\
&\le \frac{\varepsilon}{2} + \big|\langle H_n \xi, S\nu \rangle - \langle H_n \xi, H_n^{-1} \nu \rangle\big| \\
&\le \frac{\varepsilon}{2} + \big|\langle H_n \xi, \big(S - H_n^{-1}\big)\nu \rangle\big| \\
&\le \frac{\varepsilon}{2} + \frac{\varepsilon\, \|H_n \xi\|}{2 \sup\{\|H_m \xi\| :\ m\}} < \varepsilon.
\end{aligned}
$$

As $\nu$ and $\varepsilon$ were arbitrary, this shows that $S\zeta = \xi$. To achieve $\zeta \in \overline{S\mathcal{H}}$ we replace $\zeta$ with $Q\zeta$, where $Q$ is the orthogonal projection onto $\overline{S\mathcal{H}}$. Because $S$ is selfadjoint, we have $\overline{S\mathcal{H}} = (\ker S)^{\perp}$, so $S(I_{\mathcal{H}} - Q) = S$ and hence $SQ\zeta = S\zeta$.

Because $\{H_n^{-1}\}$ is bounded, by Proposition 12.1.13

$$
H_n^{-1}(4^k S_k^* S_k) H_n^{-1} \xrightarrow[\text{sot}]{} 4^k S S_k^* S_k S.
$$

By definition of $H_n$

$$
0 \le H_n^{-1}(4^k S_k^* S_k) H_n^{-1} \le H_n^{-1} \Big( I_{\mathcal{M}} + \sum_{k=1}^{n} 4^k S_k^* S_k \Big) H_n^{-1} = I_{\mathcal{M}}.
$$

It follows that

$$
\|S_k S\|^2 = \|S S_k^* S_k S\| \le 4^{-k}.
$$

This guarantees that the series

$$
T = \sum_{k=1}^{\infty} S_k S
$$

converges in norm. So $T \in \mathcal{M}$ and

$$T\zeta = \sum_{k=1}^{\infty} S_k S\zeta = \sum_{k=1}^{\infty} S_k \xi = \eta. \qquad \square$$

**12.5.1. Exercises.**

(12.5.1) In Proposition 12.5.6, where is $\sigma$-finiteness used?

(12.5.2) In the context of Proposition 12.5.6, show an example of $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$, abelian and without a separating vector.

(12.5.3) Verify the facts about the atomic masa stated after its definition (12.18). The only nontrivial part is that $\mathcal{A}_a \not\simeq \mathcal{A}_m$, which can be seen by looking at the existence or not of minimal projections in both algebras.

(12.5.4) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a maximal abelian von Neumann algebra (a masa). Show that if $\mathcal{A}$ has a cyclic (equivalently, separating) vector then $\mathcal{A}$ is $\sigma$-finite.

(12.5.5) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra such that $\xi \in \mathcal{H}$ is separating for $\mathcal{M}$. Show that $\mathcal{M}$ is $\sigma$-finite.

## 12.6. Normal Functionals

Since we have so many special topologies hanging around, it should not be a surprise that continuity with respect to said topologies is important. We recall from Exercise 12.1.22 that $\sum_j P_j$ always exists—as a sot-limit—for any pairwise orthogonal family of projections (Proposition 12.1.10). Being a sot-limit, if $\mathcal{M}$ is a von Neumann algebra and $\{P_j\} \subset \mathcal{M}$ are pairwise orthogonal projections, $\sum_j P_j \in \mathcal{M}$.

**12.6.1. Definition.**

*Let $\mathcal{M}, \mathcal{N}$ be von Neumann algebras. A linear map $\Psi : \mathcal{M} \to \mathcal{N}$ is* ***normal*** *if it is $\sigma$-weak to $\sigma$-weak continuous. In particular, a linear functional $\varphi \in \mathcal{M}^*$ is* ***normal*** *if it is $\sigma$-weak-continuous.*

*A positive linear functional $\varphi \in \mathcal{M}^*$ is* ***$\sigma$-additive*** *if whenever $\{P_j\} \subset \mathcal{M}$ is a pairwise orthogonal family of projections, $\varphi\Big(\sum_j P_j\Big) = \sum_j \varphi(P_j)$.*

*A* ***support projection*** *for $\varphi$ is a projection $F \in \mathcal{M}$ such that $\varphi(FT) = \varphi(T)$ for all $T \in \mathcal{M}$, and for any projection $P \in \mathcal{M}$, $\varphi(P) = 0$ implies $P \leq I_{\mathcal{M}} - F$.*

The adjective "normal" has already been used a lot in this text, applied to linear operators in $\mathcal{B}(\mathcal{H})$. Now we are applying it to linear functionals. There is no relation whatsoever between the two meanings, but unfortunately both are entirely standard.

Historically, normality has been often introduced for positive maps as the property of preserving suprema of monotone nets of selfadjoints. The equivalence will be addressed in (ii) of Proposition 12.6.11.

A first comment about support projections is that the definition is not claiming that they exist. And the restriction to $\sigma$-additive functionals is not capricious, as for non-normal positive functionals a support projection may not exist. Let $\rho : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ the quotient map onto the Calkin algebra, and let $\tilde{\varphi}$ be a state on $\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$. Then $\varphi = \tilde{\varphi} \circ \rho$ is a state on $\mathcal{B}(\mathcal{H})$. If $F$ is a support projection for $\varphi$, let $P \leq F$ be a finite-rank projection. Then $\varphi(P) = 0$ but $P \leq F$, showing that no support projection for $F$ can exist.

Because $\varphi$ in Definition 12.6.1 is required to be positive, if $F_\varphi$ is the support projection of $\varphi$ then $\varphi(TF_\varphi) = \varphi(T)$ for all $T \in \mathcal{M}$ (Exercise 12.6.1).

***12.6.2. Lemma.***

*Let $\varphi \in \mathcal{M}^*$ be $\sigma$-additive. Then a support projection $F_\varphi \in \mathcal{M}$ exists for $\varphi$, and it is unique. And $\varphi$ is faithful on $F_\varphi \mathcal{M} F_\varphi$.*

*Proof.* The hypothesis includes the fact that $\varphi \geq 0$. If $\varphi(P) > 0$ for all projections $P \in \mathcal{M}$, we can take $F_\varphi = I_\mathcal{M}$. Otherwise, let $\{P_j\}$ be a pairwise orthogonal family of projections in $\mathcal{M}$, maximal for the property that $\varphi(P_j) = 0$ for all $j$ (Exercise 12.6.3). We have $\sum_j P_j \in \mathcal{M}$. Let $F_\varphi = I_\mathcal{M} - \sum_j P_j$. By the $\sigma$-additivity

$$\varphi(I_\mathcal{M} - F_\varphi) = \varphi\Big(\sum_j P_j\Big) = \sum_j \varphi(P_j) = 0.$$

Then for any $T \in \mathcal{M}$ we have, using Cauchy–Schwarz,

$$\begin{aligned} |\varphi(T) - \varphi(F_\varphi T)| &= |\varphi((I_\mathcal{M} - F_\varphi)T)| = |\varphi((I_\mathcal{M} - F_\varphi)T\, I_\mathcal{M})| \\ &\leq \varphi(I_\mathcal{M})^{1/2} \varphi((I_\mathcal{M} - F_\varphi)TT^*(I_\mathcal{M} - F_\varphi))^{1/2} \\ &\leq \varphi(I_\mathcal{M})^{1/2} \|T\|\, \varphi(I_\mathcal{M} - F_\varphi)^{1/2} = 0, \end{aligned}$$

showing that $\varphi(T) = \varphi(F_\varphi T)$, and hence $\varphi(T) = \varphi(TF_\varphi)$ by Exercise 12.6.1.

Finally, suppose that $P \in \mathcal{M}$ and $\varphi(P) = 0$. Then $F_\varphi P F_\varphi \in \mathcal{M}^+$. If it is nonzero, by Corollary 12.4.17 there exist $\lambda > 0$ and a nonzero projection $Q$ with $QF_\varphi PF_\varphi Q \geq \lambda Q$. Being a spectral projection of $F_\varphi PF_\varphi$, $Q$ commutes with it. Then

$$\varphi(Q) \leq \lambda^{-1} \varphi(QF_\varphi PF_\varphi Q) = \lambda^{-1} \varphi(QF_\varphi PF_\varphi)$$

$$\leq \varphi(F_\varphi P F_\varphi)^{1/2} \varphi(Q F_\varphi Q)^{1/2} = 0,$$

since $\varphi(FPF) = \varphi(P) = 0$. So $\varphi(Q) = 0$ and

$$(I_\mathcal{M} - F_\varphi)Q(I_\mathcal{M} - F_\varphi) \leq \lambda^{-1}(I_\mathcal{M} - F)QF_\varphi PF_\varphi(I_\mathcal{M} - F) = 0,$$

so $Q \leq F_\varphi$. This would make $Q$ pairwise orthogonal with all $P_j$, contradicting the maximality. It follows that $F_\varphi P F_\varphi = 0$, so $P \leq I_\mathcal{M} - F_\varphi$ and $F_\varphi$ is a support projection for $\varphi$.

It remains to check the uniqueness. Suppose that $F_1, F_2$ are support projections for $\varphi$. We have $\varphi(F_1) = \varphi(F_1\, I_\mathcal{M}) = \varphi(I_\mathcal{M})$. Then $\varphi(I_\mathcal{M} - F_1) = 0$; as $F_2$ is a support projection, $I_\mathcal{M} - F_1 \leq I_\mathcal{M} - F_2$. Reversing the roles we get the reverse inequality, and so $F_1 = F_2$.

We leave the faithfulness as an exercise (Exercise 12.6.2). ☐

Here is the analog of Proposition 12.1.2.

***12.6.3. Proposition.***

*Let $\varphi : \mathcal{M} \to \mathbb{C}$ be linear, where $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is a $\sigma$-weakly closed subspace. The following statements are equivalent:*

*(i) $\varphi$ is normal;*

*(ii) there exists $S \in \mathcal{T}(\mathcal{H})$ such that $\varphi(T) = \mathrm{Tr}(ST)$ for all $T \in \mathcal{M}$;*

*(iii) $\varphi$ is $\sigma$-sot-continuous;*

*(iv) there exist sequences $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}$ of pairwise orthogonal vectors such that $\sum_n \|\xi_n\|^2 + \|\eta_n\|^2 < \infty$, and*

$$\varphi(T) = \sum_{k=1}^\infty \langle T\xi_k, \eta_k\rangle, \qquad T \in \mathcal{B}(\mathcal{H}) \tag{12.19}$$

*(v) there exist sequences $\{\xi_n\}, \{\eta_n\} \subset \mathcal{H}$ such that $\sum_n \|\xi_n\|^2 + \|\eta_n\|^2 < \infty$, and*

$$\varphi(T) = \sum_{k=1}^\infty \langle T\xi_k, \eta_k\rangle, \qquad T \in \mathcal{B}(\mathcal{H}) \tag{12.20}$$

*(vi) $\varphi$ is wot-continuous on bounded subsets of $\mathcal{M}$;*

*(vii) $\varphi$ is sot-continuous on bounded subsets of $\mathcal{M}$;*

*Proof.* (i) $\implies$ (ii) This is Proposition 7.2.10, after extending $\varphi$ to a $\sigma$-weakly continuous linear functional on $\mathcal{B}(\mathcal{H})$ via Corollary 5.7.24.

(ii) $\iff$ (i) By definition.

(ii) $\Longrightarrow$ (iv) By Theorem 10.6.26 there exist orthonormal sequences $\{\xi_n\}$ and $\{\eta_n\}$ in $\mathcal{H}$, and non-negative numbers $\{\sigma_n\}$ such that

$$S = \sum_n \sigma_n \xi_n \eta_n^*.$$

Then, after extending $\{\eta_n\}$ to an orthonormal basis,

$$\begin{aligned}\varphi(T) = \operatorname{Tr}(ST) = \operatorname{Tr}(TS) &= \sum_n \langle TS\eta_n, \eta_n\rangle \\ &= \sum_n \sigma_n \langle T\xi_n, \eta_n\rangle = \sum_n \langle T\sigma_n^{1/2}\xi_n, \sigma_n^{1/2}\eta_n\rangle.\end{aligned}$$

The sequences $\{\sigma_n^{1/2}\xi_n\}$ and $\{\sigma_n^{1/2}\eta_n\}$ are (each) pairwise orthogonal and

$$\sum_n \left\|\sigma_n^{1/2}\xi_n\right\|^2 = \sum_n \sigma_n \|\xi_n\|^2 = \sum_n \sigma_n \underset{(10.24)}{=} \operatorname{Tr}(|S|) < \infty$$

and the same estimate works for the other sequence.

(iv) $\Longrightarrow$ (v) Trivial.

(v) $\Longrightarrow$ (vi) Suppose that $\{T_j\} \subset \mathcal{H}$ with $\|T_j\| \le c$ for all $j$, and $T_j \xrightarrow{\text{wot}} T$. Then $\|T\| \le c$ also, since $\|T\| = \sup\{|\langle T\xi, \eta\rangle| : \|\xi\| = \|\eta\| = 1\}$. Fix $\varepsilon > 0$ and choose $n_0$ such that

$$\sum_{n>n_0} \|\xi_n\|^2 < \varepsilon, \qquad \sum_{n>n_0} \|\eta_n\|^2 < \varepsilon.$$

Then

$$\begin{aligned}|\varphi(T_j - T)| &\le \sum_{k=1}^{n_0} |\langle (T_j - T)\xi_k, \eta_k\rangle| + \sum_{k>n_0} |\langle (T_j - T)\xi_k, \eta_k\rangle| \\ &\le \sum_{k=1}^{n_0} |\langle (T_j - T)\xi_k, \eta_k\rangle| + 2c \sum_{k>n_0} \|\xi_k\|\,\|\eta_k\| \\ &\le \sum_{k=1}^{n_0} |\langle (T_j - T)\xi_k, \eta_k\rangle| + c \sum_{n>n_0} \|\xi_n\|^2 + \|\eta_n\|^2 \\ &\le \sum_{k=1}^{n_0} |\langle (T_j - T)\xi_k, \eta_k\rangle| + 2c\,\varepsilon.\end{aligned}$$

Hence $\limsup_j |\varphi(T_j - T)| \le 2c\varepsilon$ and, by the Limsup Routine, the limit exists and is zero. That is, $\varphi(T_j) \to \varphi(T)$ and $\varphi$ is wot-continuous on bounded sets.

(vi) $\Longrightarrow$ (vii) Trivial, since wot is weaker than sot.

(vii) $\Longrightarrow$ (i) By hypothesis, $\overline{B_r^{\mathcal{M}}(0)} \cap \ker\varphi$ is sot-closed for all $r > 0$ (recall that $\overline{B_1^{\mathcal{M}}(0)}$ is sot-closed by Exercise 12.1.4 and scaling will not change the fact). Being convex, it is wot-closed by Corollary 12.1.3. By Lemma 12.1.21

we get that $\overline{B_r^{\mathcal{M}}(0)} \cap \ker\varphi$ is $\sigma$-weak closed for all $r > 0$; and then by Krein–Šmulian (Theorem 7.2.25), $\ker\varphi$ is $\sigma$-weak closed, which implies that $\varphi$ is $\sigma$-weak continuous via Proposition 5.5.9.

(iii) $\iff$ (vii) This follows directly from Lemma 12.1.21. □

A remarkable consequence of Proposition 12.6.3 is that the normal functionals of any von Neumann algebra are precisely the restrictions of the normal functionals of the environment $\mathcal{B}(\mathcal{H})$.

We will now focus on what we can say about normal states. A first result is that normal linear functionals on a von Neumann algebra admit a polar decomposition; this complements the Jordan Decomposition we saw in Proposition 11.5.16. We use the notation $V\omega$ to mean the linear functional $X \longmapsto \omega(VX)$.

***12.6.4. Proposition. (Polar Decomposition for Normal Functionals)***

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ normal. Then there exist a unique partial isometry $V \in \mathcal{M}$ and $\omega \in \mathcal{M}^*$, positive and normal, such that $\varphi = V\omega$ and $VV^* = F_\omega$.*

*Proof.* Suppose, without loss of generality, that $\|\varphi\| = 1$. Consider the set

$$\mathcal{R} = \{T \in \mathcal{M} : \varphi(T) = 1,\ \|T\| \leq 1\}.$$

This set is nonempty. Indeed, since $\|\varphi\| = 1$ there exists $\{T_n\} \subset \mathcal{M}$ with $\|T_n\| = 1$ and $|\varphi(T_n)| \to 1$. Multiplying each $T_n$ by an appropriate complex number of absolute value 1, we may assume that $\varphi(T_n) \to 1$. By Proposition 12.1.8 there is a wot-convergent subnet $\{T_{n_j}\}$, with $T_{n_j} \xrightarrow{\text{wot}} T$. We have $\|T\| \leq 1$ and, by Proposition 12.6.3, the normality of $\varphi$ implies $\varphi(T_{n_j}) \to \varphi(T)$. So $T \in \mathcal{R}$.

The set $\mathcal{R}$ is convex, and it is $\sigma$-weak-compact by the normality of $\varphi$ and Proposition 12.1.8. By Krein–Milman (Theorem 7.5.11) there exists an extreme point $U \in \mathcal{R}$. Necessarily $\|U\| = 1$, and this implies that $U$ is extreme in $\overline{B_1^{\mathcal{M}}(0)}$. Indeed, if $U = tX + (1-t)Y$ with $\|X\| \leq 1$, $\|Y\| \leq 1$, and $\varphi(U) = 1$, then $1 = t\varphi(X) + (1-t)\varphi(Y)$; as $\|\varphi\| = 1$,

$$1 = \varphi(U) = t\varphi(X) + (1-t)\varphi(Y) = t\mathrm{Re}\,\varphi(X) + (1-t)\mathrm{Re}\,\varphi(Y) \leq 1.$$

Since $|\varphi(X)| \leq 1$ and $|\varphi(Y)| \leq 1$, we deduce that $\mathrm{Re}\,\varphi(X) = \mathrm{Re}\,\varphi(Y) = 1$, and so $\varphi(X) = \varphi(Y) = 1$. Thus $X, Y \in \mathcal{R}$; and $U$ is extreme in $\mathcal{R}$, so $X = Y = U$, and therefore $U$ is extreme in the closed unit ball. Now Proposition 11.3.13 tells us that $U$ is a partial isometry.

We have $\|\varphi U\| \leq \|U\|\,\|\varphi\| = 1$; we also have $(\varphi U)(I_{\mathcal{M}}) = \varphi(U) = 1$. Then $\varphi U \geq 0$ by Proposition 11.5.4. From

$$(\varphi U)(UU^*) = \varphi(UU^*U) \overset{\text{(Exercise 10.4.12)}}{=} \varphi(U) = 1$$

we deduce that $(\varphi U)(I - UU^*) = 0$, so $I_{\mathcal{M}} - UU^* \leq I_{\mathcal{M}} - F_{\varphi U}$, which is $F_{\varphi U} \leq UU^*$. Let $W = U^* F_{\varphi U}$ and $\omega = \varphi W^*$. Then $W^*W = F_{\varphi U} UU^* F_{\varphi U} = F_{\varphi U}$. For any $T \in \mathcal{M}$,

$$(\varphi U)(T) = (\varphi U)(TF_{\varphi U}) = \varphi(TF_{\varphi U}U) = \varphi(TW^*) = (\varphi W^*)(T).$$

So $\varphi U = \varphi W^* = \omega$. The next step is to show that

$$\varphi(T) = \varphi(TP), \qquad T \in \mathcal{M},$$

where $P = WW^*$, $Q = W^*W$. Fix $T \in \mathcal{M}$ and let $c = \varphi(\lambda(I_{\mathcal{M}} - P)T)$, with $\lambda \in \mathbb{T}$ such that $c \geq 0$. If $c > 0$, there exists $n \in \mathbb{N}$ with $n+c > (n^2+\|T\|^2)^{1/2}$. This gives, since $\varphi(W^*) = \varphi(U) = 1$ and $W^*P = W^*$,

$$\begin{aligned}
\|nW^*+\lambda(I_{\mathcal{M}} - P)T\| &= \left\|\big(nW^*+\lambda(I_{\mathcal{M}} - P)T\big)\big(nW^*+\lambda(I_{\mathcal{M}} - P)T\big)^*\right\|^{1/2} \\
&= \|n^2Q + (I_{\mathcal{M}} - P)TT^*(I_{\mathcal{M}} - P)\|^{1/2} \\
&\leq (n^2 + \|T\|^2)^{1/2} < n + c \\
&= \varphi\big((nW^* + \lambda(I_{\mathcal{M}} - P)T\big).
\end{aligned}$$

This contradicts the assumption that $\|\varphi\| = 1$; thus $c = 0$. We then have

$$\varphi(T) = \varphi(TP) = \varphi(TWW^*) = \varphi W^*(TW) = (\omega W)(T).$$

Next we apply the above to the normal linear functional $\tilde{\varphi}(T) = \overline{\varphi(T^*)}$ (this is linear by Proposition 10.7.5 and normal by Proposition 12.6.3 since the adjoint operation is $\sigma$-weak-continuous). So there exists a partial isometry $W \in \mathcal{M}$ and a positive normal $\omega \in \mathcal{M}^*$ such that $\tilde{\varphi} = \omega W$. Then, using that $\omega \geq 0$ and so it preserves adjoints,

$$\varphi(T) = \overline{\tilde{\varphi}(T^*)} = \overline{\omega(T^*W)} = \omega(W^*T) = (V\omega)(T),$$

where $V = W^*$.

It remains to check the uniqueness. We assume without loss of generality that $\omega(I_{\mathcal{M}}) = 1$. Suppose that $\varphi = V'\omega'$ for a positive normal $\omega'$ and a partial isometry $V'$ such that $V'V'^* = F_{\omega'}$. We have $\varphi(T) = \omega(VT) = \omega'(V'T)$. From

$$Q = W^*W = F_{\varphi U} = F_{\varphi W^*} = F_\omega,$$

we have

$$\begin{aligned}
\omega(I_{\mathcal{M}}) = \omega(F_\omega) = \omega(VV^*) &= \omega'(V'V^*) = \omega'(V'V^*F_{\omega'}) \\
&= \omega(VV^*F_{\omega'}) = \omega(F_\omega F_{\omega'}) = \omega(F_{\omega'}).
\end{aligned}$$

Thus $\omega(I_{\mathcal{M}} - F_{\omega'}) = 0$, so $I_{\mathcal{M}} - F_{\omega'} \leq I_{\mathcal{M}} - F_\omega$, so $F_\omega \leq F_{\omega'}$. Exchanging roles we deduce that $F_{\omega'} = F_\omega$.

Write $VV'^* = A + iB$ with $A, B \in F_\omega \mathcal{M} F_\omega$ selfadjoint. Then

$$\omega(A) + i\omega(B) = \omega(VV'^*) = \omega'(V'V'^*) = \omega'(F_{\omega'}) = 1.$$

Thus $\omega(B) = 0$ and $\omega(A) = 1$. Note that $\|A\| \leq \|VV'^*\| \leq 1$. If $\omega(A^-) > 0$, then $\omega(A^+) = \omega(A) + \omega(A^-) > 1$, contradicting that $\|\omega\| = 1$. Hence $\omega(A^-) = 0$, and $A^- = 0$ since $\omega$ is faithful on $F_\omega \mathcal{M} F_\omega$ (Exercise 12.6.2). Together with $\|A\| \leq 1$ we get that $A \leq F_\omega$, and so $A = F_\omega$ by the faithfulness of $\omega$. As $F_\omega$ is the identity in $F_\omega \mathcal{M} F_\omega$, we get $B = 0$ from $\|VV'^*\| \leq 1$ and Exercise 10.2.3. Therefore $VV'^* = A = F_\omega$. Now, as $V'V^* = F_\omega = F_{\omega'}$,

$$\begin{aligned} V^* &= V^*VV^* = V^*F_\omega = V^*VV'^*, \\ V'^* &= V'^*V'V'^* = V'^*F_{\omega'} = V'^*V'V^*. \end{aligned} \tag{12.21}$$

This gives

$$V^*V = V^*V(V'^*V')V^*V, \qquad V'^*V' = V'^*V'(V^*V)V'^*V'.$$

From the first equality we obtain $V^*V(I_{\mathcal{M}} - V'^*V')V^*V = 0$, and thus $(I_{\mathcal{M}} - V'^*V')V^*V = 0$, which means that $V^*V \leq V'^*V'$. In an analog way the second equality gives us $V'^*V' \leq V^*V$, so $V'^*V' = V^*V$. Going back to (12.21) with this new equality at hand, we obtain $V^* = V'^*$ and thus $V' = V$. Finally, $\omega'(T) = \omega'(VV^*T) = \omega(VV^*T) = \omega(T)$. $\square$

In a similar vein, we have an analog of Proposition 11.5.16 for normal functionals.

**12.6.5. Corollary.**

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}_*$ Hermitian. Then there exist $\varphi^+, \varphi^- \in \mathcal{M}_*^+$ with $\varphi = \varphi^+ - \varphi^-$. As a consequence, every normal linear functional on $\mathcal{M}$ is a linear combination of four normal states.*

*Proof.* Since $\varphi$ is Hermitian, in the proof of Proposition 12.6.4 we may consider only the self adjoint elements of $\mathcal{R}$, and replace $T_n$ with $\frac{\operatorname{Re} T_n}{\|\operatorname{Re} T_n\|}$ for all $n$ to show that $\mathcal{R}$ is nonempty. Then $U$ is selfadjoint. Being a selfadjoint partial isometry, by Exercise 11.2.9 there exist projections $P, Q \in \mathcal{M}$ with $U = P - Q$. As $F_{\varphi U} \leq U^2 = P + Q$, we get $F_{\varphi U}(P+Q) = F_{\varphi U} = (P+Q)F_{\varphi U}$. Multiplying separately by $P$ and $Q$, since $PQ = 0$ we get that $F_{\varphi U}$ commutes with both $P$ and $Q$, and hence it commutes with $U$. This means that $W$ is selfadjoint, and therefore $V$ is selfadjoint. By Exercise 11.2.9 there exists projections $E, F \in \mathcal{M}$ with $V = E - F$ and $EF = 0$. We have, for $T \geq 0$

$$\varphi(TV) = \overline{\varphi(VT)} = \overline{\omega(T)} = \omega(T) = \varphi(VT).$$

Iterating this to pass first to powers of $V$ and taking linear combinations and limits we get $\varphi(Tf(V)) = \varphi(f(V)T)$. So $\varphi(TE) = \varphi(ET)$ and $\varphi(TF) = \varphi(FT)$ for all $T \in \mathcal{M}$. In particular $E\varphi$ and $F\varphi$ are Hermitian. Also, if $T \geq 0$ we have

$$E\varphi(T) = \varphi(ET) = \varphi(ETE) = \varphi(ETE) - \varphi(FETE) = \omega(ETE) \geq 0.$$

So $\varphi^+ = E\varphi \geq 0$ and similarly $\varphi^- = -F\varphi \geq 0$. Then

$$\varphi = V^2\varphi = (E + F)\varphi = \varphi^+ - \varphi^-.$$

Finally, $\varphi^+$ and $\varphi^-$ are normal since $\varphi$ is, and so is multiplication by a projection. ☐

We now state a few technical lemmas to be needed soon.

***12.6.6. Lemma.***

*Let $\mathcal{M}$ be a von Neumann algebra, $\varphi \in \mathcal{M}^*$ a state, $P \in \mathcal{M}$ a projection, and $\xi \in \mathcal{H}$ with $\|\xi\| = 1$. If $\varphi(Q) \leq \langle Q\xi, \xi\rangle$ for all projections $Q \in \mathcal{M}$ such that $Q \leq P$, then*

$$\varphi(PTP) \leq \langle PTP\xi, \xi\rangle, \qquad\qquad T \in \mathcal{M}^+.$$

*Proof.* Fix $T \in \mathcal{M}^+$ and $\varepsilon > 0$. By the Spectral Theorem there exist pairwise orthogonal projections $P_1, \dots, P_n \leq P$, with $\sum_k P_k = P$, and scalars $\{\lambda_k\}$ such that

$$\Big\| PTP - \sum_{k=1}^{n} \lambda_k P_k \Big\| < \varepsilon.$$

We know from (12.15) and Corollary 12.4.15 that $P_k \in \mathcal{M}$ for all $k$. We have

$$\begin{aligned}\varphi(PTP) &\leq \varepsilon + \sum_{k=1}^{n} \lambda_k \varphi(P_k) \leq \varepsilon + \sum_{k=1}^{n} \lambda_k \langle P_k\xi, \xi\rangle \\ &= \varepsilon + \Big\langle \sum_{k=1}^{n} \lambda_k P_k\, \xi, \xi \Big\rangle \leq 2\varepsilon + \langle PTP\xi, \xi\rangle.\end{aligned}$$

As we can do this for all $\varepsilon > 0$, we obtain $\varphi(PTP) \leq \langle PTP\xi, \xi\rangle$. ☐

***12.6.7. Lemma.***

*Let $\mathcal{M}$ be a von Neumann algebra, $\varphi \in \mathcal{M}^*$ a $\sigma$-additive state, $P \in \mathcal{M}$ a projection, and $\xi \in P\mathcal{H}$ with $\|\xi\| = 1$. Then there exists a nonzero projection $P_0 \in \mathcal{M}$ such that $P_0 \leq P$ and*

$$\varphi(P_0TP_0) \leq \langle P_0TP_0\xi, \xi\rangle, \qquad\qquad T \in \mathcal{M}^+. \tag{12.22}$$

*Proof.* If $P$ satisfies (12.22), we are done. Otherwise, by Lemma 12.6.6 there exists a projection $Q \in \mathcal{M}$ with $Q \leq P$ and $\varphi(Q) > \langle Q\xi, \xi\rangle$. Using Zorn's Lemma, we construct a maximal family $\{Q_j\}$ of pairwise orthogonal subprojections of $P$ in $\mathcal{M}$ such that $\varphi(Q_j) > \langle Q_j\xi, \xi\rangle$ for all $j$. Then $Q = \sum_j Q_j$ is a subprojection of $P$ in $\mathcal{M}$ and (here we use the $\sigma$-additivity)

$$\varphi(Q) = \sum_j \varphi(Q_j) > \sum_j \langle Q_j\xi, \xi\rangle = \langle Q\xi, \xi\rangle.$$

If we had $P = Q$, it would follow that

$$\varphi(P) > \langle P\xi, \xi\rangle = 1 \geq \varphi(P),$$

a contradiction. Thus $P_0 = P - Q \in \mathcal{M}$ is nonzero. The maximality of $\{Q_j\}$ guarantees that $\varphi(Q') \leq \langle Q'\xi, \xi\rangle$ for all subprojections $Q' \leq P_0$ in $\mathcal{M}$. Then (12.22) holds by Lemma 12.6.6 . □

### *12.6.8. Lemma.*

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ a $\sigma$-additive state. Then there exist pairwise orthogonal projections $\{P_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$, and vectors $\{\xi_n\}$ such that $\|\xi_n\| \leq 1$, $P_n\xi_n = \xi_n$,*

$$\varphi(P_nTP_n) \leq \langle P_nTP_n\xi_n, \xi_n\rangle, \qquad T \in \mathcal{M}^+, \tag{12.23}$$

*and $\sum_n P_n = I_{\mathcal{M}}$.*

*Proof.* Let $F_\varphi$ be the support projection for $\varphi$. Fix $\xi \in F_\varphi\mathcal{H}$ with $\|\xi\| = 1$. By Lemma 12.6.7 there exists a projection $P_0 \in \mathcal{M}$ with $P_0 \leq F_\varphi$ and satisfying (12.22). Now use Zorn's Lemma to construct a family $\{P_j\} \subset \mathcal{M}$ of pairwise orthogonal projections $P_j \leq F_\varphi$, maximal with respect to the property that for each $j$ there exists $\xi_j \in \mathcal{H}$ with $\varphi(P_jTP_j) \leq \langle P_jTP_j\xi_j, \xi_j\rangle$ for all $T \in \mathcal{M}^+$. Because $P_j \leq F_\varphi$ we have $\varphi(P_j) > 0$ (as $\varphi$ is faithful on $F_\varphi\mathcal{M}F_\varphi$ by Lemma 12.6.2). Then, for any finite $R \subset J$,

$$\sum_{j\in R} \varphi(P_j) = \varphi\Big(\sum_{j\in R} P_j\Big) \leq 1,$$

forcing the family $\{P_j\}$ to be countable. If $I_{\mathcal{M}} - \sum_j P_j \neq 0$, by Lemma 12.6.7 there exists a projection $P' \in \mathcal{M}$ with $P' \leq I_{\mathcal{M}} - \sum_j P_j$, and $\eta \in \mathcal{H}$ such that $\varphi(P'TP') \leq \langle P'TP'\eta, \eta\rangle$ for all $T \in \mathcal{M}^+$. As this contradicts the maximality of $\{P_j\}$, it follows that $\sum_j P_j = I_{\mathcal{M}}$. Finally, if needed we replace $\xi_n$ with $P_n\xi_n$. □

**12.6.9. Proposition.**

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ a $\sigma$-additive state. Then $\varphi$ is normal.*

*Proof.* By Lemma 12.6.8 there exist unit vectors $\{\xi_n\} \subset \mathcal{H}$ and pairwise orthogonal projections $\{P_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$ that satisfy (12.23) and also $\sum_n P_n = I_{\mathcal{M}}$. Fix $\varepsilon > 0$, and $n_0$ such that $\sum_{n>n_0} \varphi(P_n) < \varepsilon^2/4$. Let

$$N_\varphi = \{T \in \mathcal{B}(\mathcal{H}) : \ \|TP_n\xi_n\| < \varepsilon/(2n_0), \ n = 1, \ldots, n_0\}.$$

Then $N_\varphi$ is a sot-neighbourhood of 0. For $T \in N_\varphi \cap \mathcal{M}$ and $\|T\| \leq 1$,

$$\begin{aligned}
|\varphi(T)|^2 \leq \varphi(T^*T) &= \varphi\Big(\sum_{n=1}^{n_0} T^*TP_n\Big) + \varphi\Big(T^*T\sum_{n>n_0} P_n\Big) \\
&\leq \sum_{n=1}^{n_0} \varphi(|T|\,|T|P_n) + \varphi((T^*T)^2)^{1/2}\varphi\Big(\sum_{n>n_0} P_n\Big)^{1/2} \\
&\leq \sum_{n=1}^{n_0} \varphi(|T|^2)^{1/2}\varphi(P_n|T|^2P_n)^{1/2} + \frac{\varepsilon}{2} \\
&\leq \sum_{n=1}^{n_0} \langle P_nT^*TP_n\xi_n, \xi_n\rangle^{1/2} + \frac{\varepsilon}{2} \\
&= \sum_{n=1}^{n_0} \|TP_n\xi_n\| + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Thus $\varphi$ is sot-continuous at 0 on $\overline{B_1^{\mathcal{M}}(0)}$. Translating and scaling, we get that $\varphi$ is sot-continuous on bounded sets. By Proposition 12.6.3, $\varphi$ is normal. $\square$

**12.6.10. Theorem. (Sakai's Radon–Nikodym)**

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi, \psi \in \mathcal{M}^*$ positive linear functionals such that $\varphi \leq \psi$. Then there exists $A \in \mathcal{M}$, positive, with $\varphi(T) = \psi(ATA)$ for all $T \in \mathcal{M}$.*

*Proof.* Let $F_\psi$ be the support projection for $\psi$. By working on $F_\psi\mathcal{M}F_\psi$ and by Lemma 12.6.2, we may assume without loss of generality that $\psi$ is faithful. Consider the GNS representation $\pi_\psi : \mathcal{M} \to \mathcal{B}(\mathcal{H}_\psi)$; since $\psi$ is faithful, $\pi_\psi$ is faithful (Exercise 11.6.6). We have $\psi(T) = \langle \pi_\psi(T)\xi_\psi, \xi_\psi\rangle$, and $\xi_\psi$ is cyclic and separating (the latter, by the faithfulness) for $\pi_\psi(\mathcal{M})$. Let

$$[\pi_\psi(T)\xi_\psi, \pi_\psi(S)\xi_\psi] = \varphi(S^*T).$$

This well-defined because $\psi$ is faithful and $\xi_\psi$ is separating, and it is sesquilinear. It is positive because $\varphi$ is positive. Also,

$$\begin{aligned}|\,[\pi_\psi(T)\xi_\psi, \pi_\psi(S)\xi_\psi]\,| = |\varphi(S^*T)| &\le \varphi(T^*T)^{1/2}\varphi(S^*S)^{1/2}\\ &\le \psi(T^*T)^{1/2}\psi(S^*S)^{1/2} = \|\pi_\psi(T)\xi_\psi\|\,\|\pi_\psi(S)\xi_\psi\|.\end{aligned}$$

So the form is bounded. By Proposition 10.1.5 there exists $\tilde{Y} \in \mathcal{B}(\mathcal{H}_\psi)^+$ such that

$$[\pi_\psi(T)\xi_\psi, \pi_\psi(S)\xi_\psi] = \langle \tilde{Y}\pi_\psi(T)\xi_\psi, \pi_\psi(S)\xi_\psi\rangle. \tag{12.24}$$

We have that $\tilde{Y} \in \pi_\psi(\mathcal{M})'$. Indeed, for $R, S, T \in \mathcal{M}$,

$$\begin{aligned}\langle \tilde{Y}\pi_\psi(T)\pi_\psi(S)\xi_\psi, \pi_\psi(R)\xi_\psi\rangle &= \varphi(R^*(TS)) = \varphi((T^*R)^*S)\\ &= \langle \tilde{Y}\pi_\psi(S)\xi_\psi, \pi_\psi(T^*R)\xi_\psi\rangle\\ &= \langle \pi_\psi(T)\tilde{Y}\pi_\psi(S)\xi_\psi, \pi_\psi(R)\xi_\psi\rangle.\end{aligned}$$

As $R, S$ are arbitrary and $\xi_\psi$ is cyclic for $\psi$, this shows that $\tilde{Y}\pi_\psi(T) = \pi_\psi(T)\tilde{Y}$ for all $T \in \mathcal{M}$, so $\tilde{Y} \in \pi_\psi(\mathcal{M})'$. Now define $\gamma : \pi_\psi(\mathcal{M})' \to \mathbb{C}$ to be

$$\gamma(\tilde{Z}) = \langle \tilde{Y}^{1/2}\tilde{Z}\xi_\psi, \xi_\psi\rangle.$$

We write $\tilde{\psi}$ for the positive linear functional $\tilde{\psi}(\tilde{Z}) = \langle \tilde{Z}\xi_\psi, \xi_\psi\rangle$. That is, $\gamma = \tilde{Y}^{1/2}\,\tilde{\psi}$. Let $\gamma = V|\gamma|$ be the polar decomposition as in Proposition 12.6.4, so $V \in \pi_\psi(\mathcal{M})'$. We have, for $\tilde{Z} \in \pi_\psi(\mathcal{M})'$, positive, and using the inequality from Exercise 11.5.4,

$$|\gamma|(\tilde{Z}) = \gamma(V^*\tilde{Z}) = \tilde{\psi}(\tilde{Y}^{1/2}V^*\tilde{Z}) \le \|\tilde{Y}^{1/2}V^*\|\,\tilde{\psi}(\tilde{Z}).$$

By the first part of the proof, applied to $|\gamma| \le \tilde{\psi}$ on $\pi_\psi(\mathcal{M})'$ (as $\xi_\psi$ is also cyclic and separating for $\pi_\psi(\mathcal{M})'$ by Proposition 12.5.2) there exists $\tilde{X} \in \pi_\varphi(\mathcal{M})'' = \pi_\varphi(\mathcal{M})$ (this equality due to Exercise 12.6.6), positive, such that $|\gamma|(\tilde{Z}) = \tilde{\psi}(\tilde{X}\tilde{Z})$. We can now write, for all $\tilde{Z} \in \pi_\psi(\mathcal{M})'$,

$$\langle \tilde{X}\tilde{Z}\xi_\psi, \xi_\psi\rangle = \tilde{\psi}(\tilde{X}\tilde{Z}) = |\gamma|(\tilde{Z}) = \tilde{\psi}(\tilde{Y}^{1/2}V^*\tilde{Z}) = \langle \tilde{Y}^{1/2}V^*\tilde{Z}\xi_\psi, \xi_\psi\rangle.$$

Writing the above equality as $\langle \tilde{Z}\xi_\psi, \tilde{X}\xi_\psi\rangle = \langle \tilde{Z}\xi_\psi, \tilde{V}Y^{1/2}\xi_\psi\rangle$, we get $\tilde{X}\xi_\psi = V\tilde{Y}^{1/2}\xi_\psi$. Then

$$V^*\tilde{X}\xi_\psi = V^*V\,\tilde{Y}^{1/2}\xi_\psi.$$

We also have

$$\langle \tilde{Y}^{1/2}V^*V\tilde{Z}\xi_\psi, \xi_\psi\rangle = |\gamma|(V\tilde{Z}) = \gamma(\tilde{Z}) = \langle \tilde{Y}^{1/2}\tilde{Z}\xi_\psi, \xi_\psi\rangle.$$

Hence $V^*V\tilde{Y}^{1/2}\xi_\psi = \tilde{Y}^{1/2}\xi_\psi$. Using (12.24) and with $A = \pi_\psi^{-1}(\tilde{X}) \in \mathcal{M}_+$,

$$\begin{aligned}\varphi(T) = [\pi_\psi(T)\xi_\psi, \xi_\psi] &= \langle \tilde{Y}\pi_\psi(T)\xi_\psi, \xi_\psi\rangle = \langle \pi_\psi(T)\tilde{Y}^{1/2}\xi_\psi, \tilde{Y}^{1/2}\xi_\psi\rangle\\ &= \langle \pi_\psi(T)\tilde{Y}^{1/2}\xi_\psi, V^*V\tilde{Y}^{1/2}\xi_\psi\rangle = \langle \pi_\psi(T)\tilde{Y}^{1/2}\xi_\psi, V^*\tilde{X}\xi_\psi\rangle\\ &= \langle \pi_\psi(T)V\tilde{Y}^{1/2}\xi_\psi, \tilde{X}\xi_\psi\rangle = \langle \pi_\psi(T)\tilde{X}\xi_\psi, \tilde{X}\xi_\psi\rangle\end{aligned}$$

$$= \psi(ATA). \qquad \square$$

The last few results together allow us to prove the following version of Proposition 12.6.3 in the positive case. The implication (i) $\implies$ (iv) below is the nontrivial part. We know how to extend linear functionals while preserving the norm (Hahn–Banach), and when we do it to positive unital functionals we preserve positivity (Theorem 11.5.7). We also know how to extend normal functionals while preserving normality (Corollary 5.7.24). But we don't know how to combine both results! From Proposition 12.6.3, every normal state on $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ extends to all of $\mathcal{B}(\mathcal{H})$ automatically, since the functional $T \longmapsto \operatorname{Tr}(ST)$ makes sense in all of $\mathcal{B}(\mathcal{H})$. The problem is that the fact that $T \longmapsto \operatorname{Tr}(ST)$ is positive on $\mathcal{M}$ does **not** imply that $S$ is positive, so the extension is not necessarily positive. For instance let $\mathcal{M} = \mathbb{C}\, I_2 \subset M_2(\mathbb{C})$, and $\varphi(\lambda I_2) = \lambda$. Then $\varphi(\lambda I_2) = \operatorname{Tr}(S(\lambda I_2)) = \lambda \operatorname{Tr}(S)$ for any $S$ with $\operatorname{Tr}(S) = 1$, positive or not. There is no obvious way to choose a positive $S$, though we will see that it is always possible. We will use Sakai's Radon Nikodym Theorem for that.

**12.6.11. Proposition.**

*Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ positive. The following statements are equivalent:*

*(i)* $\varphi$ *is normal;*

*(ii)* *for any bounded monotone increasing net* $\{T_j\} \subset \mathcal{M}^{\rm sa}$,
$$\varphi(\sup_j T_j) = \sup_j \varphi(T_j);$$

*(iii)* $\varphi$ *is* $\sigma$*-additive;*

*(iv)* *there exists* $S \in \mathcal{T}(\mathcal{H})$, *positive, such that* $\varphi(T) = \operatorname{Tr}(ST)$ *for all* $T \in \mathcal{M}$;

*(v)* *there exists* $\{\xi_n\} \subset \mathcal{H}$, *with pairwise orthogonal elements, such that* $\sum_n \|\xi_n\|^2 < \infty$ *and*
$$\varphi(T) = \sum_{k=1}^\infty \langle T\xi_k, \xi_k\rangle, \qquad T \in \mathcal{B}(\mathcal{H}) \tag{12.25}$$

*(vi)* *there exists a sequence* $\{\xi_n\} \subset \mathcal{H}$ *such that* $\sum_n \|\xi_n\|^2 < \infty$ *and*
$$\varphi(T) = \sum_{k=1}^\infty \langle T\xi_k, \xi_k\rangle, \qquad T \in \mathcal{B}(\mathcal{H}) \tag{12.26}$$

*Proof.* (i) $\implies$ (ii) The net $\{T_j\}$ is bounded and converges sot to some self-adjoint $T$ by Proposition 12.1.10. As $\varphi$ is sot-continuous on bounded sets by Proposition 12.6.3 and the net $\{\varphi(T_j)\}$ is monotone by positivity,

$$\varphi(\sup_j T_j) = \varphi(\lim_{\text{sot}} T_j) = \lim_j \varphi(T_j) = \sup_j \varphi(T_j).$$

(ii) $\implies$ (iii) $\sum_j P_j$ is a sot-limit of an increasing net of projections.

(iii) $\implies$ (i) This is Proposition 12.6.9.

(i) $\implies$ (iv) By Proposition 12.6.3 there exists $S_0 \in \mathcal{T}(\mathcal{H})$ with $\varphi(T) = \operatorname{Tr}(S_0 T)$ for all $T \in \mathcal{B}(\mathcal{H})$. We can write $S_0 = AB$, with $A, B$ Hilbert–Schmidt and $B$ positive (just take $S_0 = V|S_0|$ the polar decomposition, and $A = V|S_0|^{1/2}$, $B = |S_0|^{1/2}$). Let $R = A + B$; then $RR^* \in \mathcal{T}(\mathcal{H})$. Let $\psi$ be the positive normal linear functional $\psi(T) = \frac{1}{4}\operatorname{Tr}(RR^*T)$. From

$$(A+B)(A+B)^* - (A-B)(A-B)^* = 4\operatorname{Re} AB = 4\operatorname{Re} S_0$$

we get, for $T \in \mathcal{M}^+$,

$$\begin{aligned}
\varphi(T) = \operatorname{Re}\varphi(T) = \operatorname{Re}\,\operatorname{Tr}(S_0T) &= \operatorname{Tr}((\operatorname{Re} S_0)T) = \operatorname{Tr}\left(T^{1/2}(\operatorname{Re} S_0)\,T^{1/2}\right)\\
&= \frac{1}{4}\operatorname{Tr}\left(T^{1/2}((A+B)(A+B)^* - (A-B)(A-B)^*)T^{1/2}\right)\\
&= \frac{1}{4}\operatorname{Tr}\left(((A+B)(A+B)^*T\right) - \frac{1}{4}\operatorname{Tr}\left((A-B)(A-B)^*)T\right)\\
&= \psi(T) - \frac{1}{4}\operatorname{Tr}\left((A-B)(A-B)^*T\right).
\end{aligned}$$

So $\varphi \le \psi$. We can write $\psi(T) = \operatorname{Tr}(R_0TR_0)$, where $R_0 = \frac{1}{2}(RR^*)^{1/2}$. By Sakai's Radon Nikodym Theorem there exists $A_0 \in \mathcal{M}$, positive, with $\varphi(T) = \operatorname{Tr}(R_0A_0TA_0R_0) = \operatorname{Tr}(A_0R_0^2A_0T)$. As $S = A_0R_0^2A_0 \ge 0$, $\varphi$ admits a positive normal extension $T \longmapsto \operatorname{Tr}(ST)$ to $\mathcal{B}(\mathcal{H})$.

(iv) $\implies$ (v) By the Spectral Theorem for Compact Operators, as in Remark 10.6.13, we can write $S = \sum_r \lambda_r P_r$ with $\lambda_r \ge 0$ for all $r$, $P_r$ pairwise orthogonal rank-one projections, and

$$\sum_r \lambda_r = \operatorname{Tr}(S)$$

by Proposition 10.7.5. Let $\{\xi_r'\} \subset \mathcal{H}$ be unit vectors with $P_r\xi_r' = \xi_r'$ for all $r$. Then $\{\xi_r'\}$ is orthonormal. If we put $\xi_r = \lambda_r^{1/2}\xi_r'$, then $\{\xi_r\}$ are pairwise orthogonal,

$$\varphi(T) = \operatorname{Tr}(ST) = \operatorname{Tr}(S^{1/2}TS^{1/2}) = \sum_r \langle TS^{1/2}\xi_r', S^{1/2}\xi_r'\rangle = \sum_r \langle T\xi_r, \xi_r\rangle,$$

and $\sum_r \|\xi_r\|^2 = \sum_r \lambda_r = \operatorname{Tr}(S) < \infty$.

(v) $\implies$ (vi) Trivial.

(vi) $\implies$ (i) This was proven in Proposition 12.6.3. $\square$

The proof of the implication (i) $\implies$ (iv) above is convoluted: it uses a nice trick and Sakai's Radon Nikodym Theorem; this in turn depends on several previous results and—crucially—on the polar decomposition of a linear functional. The proof of these results runs several pages. There are other ways to prove that any normal linear functional on a von Neumann algebra is implemented by a trace-class operator, but none of them is too smooth. For instance in [KR97b] the proof runs for more than two pages (Theorems 7.1.8 and 7.1.9) without counting other previous results. We favour our approach because instead of writing a long ad-hoc proof, we are proving important results along the way.

One interesting feature of Proposition 12.6.11 is that normality can be characterized in ways that are intrinsic to $\mathcal{M}$ and do not depend on the environment $\mathcal{B}(\mathcal{H})$. Another major consequence is that isomorphisms of von Neumann algebras can be determined fairly explicitly (Corollary 12.6.16).

***12.6.12. Corollary.***

*Let $\mathcal{M}, \mathcal{N}$ be von Neumann algebras and $\Psi : \mathcal{M} \to \mathcal{N}$ a $*$-isomorphism. Then $\Psi$ is normal (i.e., $\sigma$-weak to $\sigma$-weak continuous).*

*Proof.* Because $\Psi$ is a $*$-isomorphism it preserves order ($X \leq Y \iff Y - X = Z^*Z \iff \Psi(Y) - \Psi(X) = \Psi(Z)^*\Psi(Z) \iff \Psi(X) \leq \Psi(Y)$), so it preserves suprema. Let $\varphi \in \mathcal{N}^*$ be a normal state. Then $\varphi$ also preserves suprema by Proposition 12.6.11 and thus $\varphi \circ \Psi$ is a state on $\mathcal{M}$ that preserves suprema. By Proposition 12.6.11 we get that $\varphi \circ \Psi$ is normal. That is, if $T_j \xrightarrow{\sigma\text{-weak}} T$ in $\mathcal{M}$, then $\varphi(\Psi(T_j - T)) \to 0$. This we can write $\varphi(\Psi(T_j) - \Psi(T)) \to 0$. As we can take $\varphi$ to be $\varphi(X) = \mathrm{Tr}(SX)$ for any $S \in \mathcal{T}(\mathcal{H})$, we have that $\Psi(T_j) \xrightarrow{\sigma\text{-weak}} \Psi(T)$. $\square$

*12.6.13. Remark*. The qualifier "isomorphism" is crucial in Corollary 12.6.12. Concretely, non-normal $*$-homomorphisms of von Neumann algebras exist—even surjective ones. For instance let $\mathcal{M} = \ell^\infty(\mathbb{N})$ and $\mathcal{N} = \mathbb{C}$. Seeing $\mathcal{M}$ as $C(\beta\mathbb{N})$ (Example 11.2.9) we can choose $t \in \beta\mathbb{N} \setminus \mathbb{N}$ (that is, a free ultrafilter), and then the map $\gamma : \mathcal{M} \to \mathcal{N}$ given by $\gamma : g \longmapsto g(t)$, is not normal. Indeed, consider the sequence $x_n = \sum_{k=1}^n e_k$. Then $x_n \nearrow 1$ in $\mathcal{M}$, but $\gamma(x_n) = 0$ for all $n$ and $\gamma(1) = 1$. We can see $\gamma$ as a non-normal irreducible representation of $\mathcal{M}$. Another natural example is obtained by considering the quotient map $q : \mathcal{B}(\mathcal{H}) \to C(\mathcal{H}) = \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ for infinite-dimensional $\mathcal{H}$. If we consider a faithful representation of $C(\mathcal{H})$ into $\mathcal{B}(\mathcal{K})$ for some Hilbert space $\mathcal{K}$, then $q : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ is a non-normal $*$-homomorphism.

**12.6.14. Corollary.**

*Let $\mathcal{M}, \mathcal{N}$ be von Neumann algebras and $\Psi : \mathcal{M} \to \mathcal{N}$ a $\sigma$-weak to $\sigma$-weak continuous $*$-homomorphism. Then $\Psi(\mathcal{M})$ is $\sigma$-weak closed in $\mathcal{N}$ (and hence a von Neumann subalgebra of $\mathcal{N}$).*

*Proof.* We have that $\ker \Psi$ is a $\sigma$-weak closed ideal of $\mathcal{M}$; then from Exercise 12.1.4 and Lemma 12.1.21 we obtain that $\ker \Psi$ is wot-closed. By Corollary 12.3.12 there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ with $\ker \Psi = Z\mathcal{M}$. Then the restriction $\Psi_0 : (I_{\mathcal{M}} - Z)\mathcal{M} \to \mathcal{N}$ is an injective normal $*$-homomorphism with the same range as $\Psi$. So we may assume without loss of generality that $\Psi$ is injective. In particular, it is isometric (Proposition 11.4.9) so it maps $\overline{B_1^{\mathcal{M}}(0)}$ onto $\overline{B_1^{\Psi(\mathcal{M})}}$. Because $\overline{B_1^{\mathcal{M}}(0)}$ is $\sigma$-weak compact (Corollary 12.3.20), so is its image. Then Krein–Šmulian guarantees that $\Psi(\mathcal{M})$ is $\sigma$-weak-closed. □

One remarkable consequence of Corollary 12.6.12 is that $*$-isomorphisms of von Neumann algebras cannot get too wild. We will need a couple definitions.

**12.6.15. Definition.**

*Let $\mathcal{A}$ be a $C^*$-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. An* ***amplification*** *of $\pi$ is a representation $\tilde{\pi} : \mathcal{A} \to \bigoplus_{k \in K} \mathcal{H}$ for some set $K$, where $\tilde{\pi}(a)(\{\xi_k\}_k) = \{\pi(a)\xi_k\}_k$. A* ***reduction*** *of $\pi$ is a representation $\pi_P(a) = P\pi(a)|_{P\mathcal{H}}$, where $P$ is a projection in $\pi(\mathcal{A})'$.*

It turns out that the only way to construct isomorphisms of von Neumann algebras is by doing amplifications, reductions, and conjugating by unitaries:

**12.6.16. Corollary.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be von Neumann algebras and $\rho : \mathcal{M} \to \mathcal{N}$ a $*$-isomorphism. Then there exists an amplification $\tilde{\pi} : \mathcal{M} \to \mathcal{H}^{(m)}$ of the identity representation $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{H})$, a projection $P \in \tilde{\pi}(\mathcal{M})'$, and a unitary $U : \mathcal{K} \to P\mathcal{H}^{(m)}$, such that*

$$\rho(T) = U^* P \tilde{\pi}(T) P U, \qquad T \in \mathcal{M}. \tag{12.27}$$

*Proof.* We have that $\rho$ is normal by Corollary 12.6.12. By Propositions 11.6.3 and 11.6.8 there exists $K \subset S(\mathcal{M})$ such that $\rho = \bigoplus_{\varphi \in K} \pi_\varphi$, where $\mathcal{K} = \bigoplus_{\varphi \in K} \mathcal{H}_\varphi$. The normality of $\rho$ implies that each $\pi_\varphi$ is normal, which in turn implies that each $\varphi$ is normal (Exercise 12.6.5). Fix $\varphi \in K$. By applying Proposition 12.6.11 to $\varphi$ there exist pairwise orthogonal vectors $\{\xi_n\}_{n \in N_\varphi} \subset$

$\mathcal{H}$ such that
$$\varphi(T) = \sum_{n\in N_\varphi} \langle T\xi_n, \xi_n\rangle.$$
Let $\tilde{\mathcal{H}}_\varphi = \bigoplus_{n\in N_\varphi} \mathcal{H}$ and $\pi^\varphi : \mathcal{M} \to \mathcal{B}(\tilde{\mathcal{H}}_\varphi)$ the representation
$$\pi^\varphi(T)\Big(\bigoplus_{n\in N_\varphi} \eta_n\Big) = \bigoplus_{n\in N_\varphi} T\eta_n.$$
Let $\tilde{\xi} = \bigoplus_{n\in N_\varphi} \xi_n$. Then $\langle \pi^\varphi(T)\tilde{\xi}, \tilde{\xi}\rangle = \sum_{n\in N_\varphi} \langle T\xi_n, \xi_n\rangle = \varphi(T)$. Let $P_\varphi$ be the orthogonal projection onto $\overline{\pi^\varphi(\mathcal{M})\tilde{\xi}}$. By Exercise 11.6.8 we have $P_\varphi \in \pi^\varphi(\mathcal{M})'$, and then $P_\varphi\pi^\varphi$ is a cyclic representation of $\mathcal{M}$ with $\langle P_\varphi\pi^\varphi(T)\tilde{\xi}, \tilde{\xi}\rangle = \varphi(T)$. Now Proposition 11.6.3 guarantees that there exists a unitary $U_\varphi : \mathcal{H}_\varphi \to P_\varphi\tilde{\mathcal{H}}_\varphi$ such that $\pi_\varphi = U_\varphi^* P_\varphi \pi^\varphi P_\varphi U_\varphi$. Let
$$U = \bigoplus_{\varphi\in K} U_\varphi, \qquad P = \bigoplus_{\varphi\in K} P_\varphi, \qquad \tilde{\pi}(T) = \bigoplus_{\varphi\in K,\ n\in N_\varphi} T.$$
Let also $\tilde{\mathcal{H}} = \bigoplus_{\varphi\in K} \tilde{\mathcal{H}}_\varphi = \bigoplus_{\varphi\in K, n\in N_\varphi} \mathcal{H}$. So $\tilde{\pi} : \mathcal{M} \to \tilde{\mathcal{H}}$ is an amplification of the identity representation, $U : \mathcal{K} \to P\tilde{\mathcal{H}}$ is a unitary, and
$$\rho(T) = \bigoplus_{\varphi\in K} \pi_\varphi = \bigoplus_{\varphi\in K} U_\varphi^* P_\varphi \pi^\varphi(T) P_\varphi U_\varphi = U^* P \tilde{\pi}(T) P U. \qquad \square$$

The notation above can be made slicker by using tensor products, that we will consider in Section 13.1; indeed, a natural way to write (12.27) is
$$\rho(T) = U^* P (T \otimes I) P U.$$
We refer to Exercise 13.4.2 and subsection 13.4.2.

The **central support** of a projection $P$ in a von Neumann algebra $\mathcal{M}$ is the smallest central projection $c(P)$ such that $P \leq c(P)$. The formal definition will be delayed until Definition 14.1.2.

***12.6.17. Corollary.***

*Let $\mathcal{M}_1 \subset \mathcal{B}(\mathcal{H}_1)$ and $\mathcal{M}_2 \subset \mathcal{B}(\mathcal{H}_2)$ be isomorphic von Neumann algebras. Then there exist a Hilbert space $\mathcal{H}$, a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, and projections $P, Q \in \mathcal{M}'$ with central support $I_\mathcal{M}$ such that $\mathcal{M}_1 \simeq \mathcal{M}Q$ and $\mathcal{M}_2 \simeq \mathcal{M}P$.*

*Proof.* Let $\rho : \mathcal{M}_1 \to \mathcal{M}_2$ be a $*$-isomorphism. We will use the notation of Corollary 12.6.16. Let $\mathcal{H} = \mathcal{H}_1^{(m)}$. We have $U\mathcal{M}_2U^* = \tilde{\pi}(\mathcal{M}_1)P$. Let $\mathcal{M} = \tilde{\pi}(\mathcal{M}_1)$. Then $\mathcal{M}_1 \simeq \mathcal{M}$ and $\mathcal{M}_2 \simeq U\mathcal{M}_2U^* = \mathcal{M}P$. It follows that $X \longmapsto XP$ realizes the isomorphism between the images of $\mathcal{M}_1$ and $\mathcal{M}_2$

in $\mathcal{M}$ (that is, between $\tilde{\pi}(\mathcal{M})$ and $U\mathcal{M}_2U^* = \mathcal{M}P$). If $Z \in \mathcal{Z}(\mathcal{M}')$ is a projection with $ZP = P$ then, as $Z \in \mathcal{Z}(\mathcal{M}') = \mathcal{Z}(\mathcal{M})$ and $X \longmapsto XP$ is an isomorphism, we get that $Z = I_{\mathcal{H}}$. That is, the central support projection of $P$ is $I_{\mathcal{H}}$. ☐

**12.6.1. Exercises.**

(12.6.1) Let $\varphi \in \mathcal{M}$ be a positive normal functional with support projection $F_\varphi$. Show that $\varphi(TF_\varphi) = \varphi(T)$ for all $T \in \mathcal{M}$.

(12.6.2) Let $\varphi$ be a state on $\mathcal{M}$ and $F$ a support projection for $\varphi$. Show that $\varphi$ is faithful when restricted to $F\mathcal{M}F$.

(12.6.3) Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$ such that there exists a projection $P \in \mathcal{M}$ with $\varphi(P) = 0$. Show that there exists a pairwise orthogonal family $\{P_j\} \subset \mathcal{M}$, maximal with respect to the property that $\varphi(P_j) = 0$ for all $j$.

(12.6.4) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be von Neumann algebras, and $U : \mathcal{H} \to \mathcal{K}$ a unitary such that $U\mathcal{M}U^* \subset \mathcal{N}$. Show that $\Gamma(T) = UTU^*$ is a $\sigma$-weak continuous $*$-monomorphism.

(12.6.5) Let $\mathcal{M}$ be a von Neumann algebra and $\varphi \in \mathcal{M}^*$. Show that $\varphi$ is normal if and only if its GNS representation $\pi_\varphi$ is normal.

(12.6.6) Let $\mathcal{M}$ be a von Neumann algebra and $\psi \in \mathcal{M}^*$ normal and faithful. Show that $\pi_\psi(\mathcal{M})'' = \pi_\psi(\mathcal{M})$. *(This is a direct consequence of Corollary 12.6.12, but it is needed earlier in the text)*

## 12.7. Preduals and the Enveloping von Neumann Algebra

**12.7.1. Unique Predual.** We know that every von Neumann algebra has a predual; this follows from Theorem 10.7.11, Proposition 12.3.19, and Corollary 7.3.8. A quick glance at Propositions 12.1.8 and 7.2.26 and Lemma 12.1.21 suggests that normal functionals could make up a predual for $\mathcal{M}$. They do.

**12.7.1. Definition.**

*Let $\mathcal{M}$ be a von Neumann algebra. We denote by $\mathcal{M}_*$ the space of all normal functionals on $\mathcal{M}$, and we call it the* ***predual*** *of $\mathcal{M}$.*

Not only do von Neumann algebras have a predual, but it is unique:

***12.7.2. Theorem.***

*Let $\mathcal{M}$ be a von Neumann algebra. Then $\mathcal{M}_*$ is a Banach subspace of $\mathcal{M}^*$, and it is the unique (up to isometric isomorphism) predual of $\mathcal{M}$.*

*Proof.* Continuity is preserved by linear combinations, so $\mathcal{M}_*$ is a vector subspace of $\mathcal{M}^*$. We leave the fact that $\mathcal{M}_*$ is norm-closed as an exercise (Exercise 12.7.1). We want to show that $\mathcal{M}_*$ satisfies the conditions in Proposition 7.2.26. Given $T \in \mathcal{M}$, fix $\varepsilon > 0$ and $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ and $\|T\xi\| > \|T\| - \varepsilon$. Further, let $\eta = \frac{T\xi}{\|T\xi\|}$. Then $\varphi(S) = \langle S\xi, \eta\rangle$ defines a normal functional on $\mathcal{M}$ with $\|\varphi\| = \|\xi\|\,\|\eta\| = 1$ (Exercises 10.7.11 and 10.7.13) and $\varphi(T) = \|T\xi\| > \|T\| - \varepsilon$. It follows that $\mathcal{M}_*$ norms $\mathcal{M}$. The unit ball of $\mathcal{M}$ is $\sigma$-weak-compact by Corollary 12.3.20, and so $\mathcal{M}_*$ is a predual for $\mathcal{M}$ by Proposition 7.2.26.

We now turn to the uniqueness. If $\mathcal{X}$ is a predual for $\mathcal{M}$, then $\mathcal{X}$ embeds isometrically into $\mathcal{X}^{**} = \mathcal{M}^*$ by Corollary 7.2.5. So we assume $\mathcal{X} \subset \mathcal{M}^*$. We use the notation $\mathcal{X}^+ = \{\varphi \in \mathcal{X} : \ \varphi \geq 0\}$. We will soon show that

$$\mathcal{M}^{\text{sa}} \ \text{ is weak}^*\text{-closed;} \tag{12.28}$$

$$T \in \mathcal{M}^+ \iff \varphi(T) \geq 0 \ \text{ for all } \varphi \in \mathcal{X}^+; \tag{12.29}$$

$$T = 0 \iff \varphi(T) = 0 \ \text{ for all } \varphi \in \mathcal{X}^+; \tag{12.30}$$

$$\mathcal{X} = \overline{\operatorname{span}}^{\|\cdot\|}\, \mathcal{X}^+. \tag{12.31}$$

Assuming (12.28)-(12.31) hold, let $\{T_j\} \subset \mathcal{M}^{\text{sa}}$ a bounded increasing net with $T_j \xrightarrow{\text{sot}} T$. As $\mathcal{M}^{\text{sa}} \cap \overline{B^{\mathcal{M}}_{\|T\|}(0)}$ is weak$^*$-compact (since $\mathcal{M}^{\text{sa}}$ is weak$^*$-closed by (12.28) and the closed unit ball is weak$^*$-compact by Corollary 12.3.20), there is a weak$^*$-convergent subnet $\{T_{j_k}\}$, say $T_{j_k} \xrightarrow{\text{weak}^*} S$; this by definition means that $\varphi(T_{j_k}) \to \varphi(S)$ for all $\varphi \in \mathcal{X}$. When $\varphi \geq 0$, it preserves order. The net of numbers $\{\varphi(T_j)\}$ is bounded and monotone—hence convergent—so it necessarily has the same limit as the subnet: $\varphi(T_j) \to \varphi(S)$. As $\varphi(T_j) \leq \varphi(T)$ for all $j$, $\varphi(S) \leq \varphi(T)$. This implies, by (12.29), that $S \leq T$; but $T = \sup_j T_j$ by Proposition 12.1.10 and from $\varphi(T_j) \leq \varphi(S)$ for all $\varphi$ we also know that $T_j \leq S$ for all $j$, so $S = T$. By Proposition 12.6.11 any $\psi \in \mathcal{X}^+$ is normal and hence, as $\mathcal{X}$ is the closed linear span of its positive elements and norm limits of normal functionals are normal functionals (Exercise 12.7.1), $\mathcal{X} \subset \mathcal{M}_*$. If the inclusion were proper, we would also have a proper inclusion $\overline{B^{\mathcal{X}}_1(0)} \subsetneq \overline{B^{\mathcal{M}_*}_1(0)}$. By Theorem 7.1.16 $\overline{B^{\mathcal{X}}_1(0)}$ is closed in the $\sigma(\mathcal{X}, \mathcal{M})$ topology, which is precisely the weak$^*$-topology when we consider $\mathcal{X}$ as a subspace of $\mathcal{M}^*$. Let $\psi \in \overline{B^{\mathcal{M}_*}_1(0)} \setminus \overline{B^{\mathcal{X}}_1(0)}$. Applying Hahn–Banach (Corollary 5.7.19) to $\{\psi\}$ and $\overline{B^{\mathcal{X}}_1(0)}$ in $\mathcal{M}^*$ with the weak$^*$-topology, via Proposition 7.2.10 there exists $T \in \mathcal{M}$ such that $\psi(T) = 1$ and $\varphi(T) = 0$ for all $\varphi \in \overline{B_1^{\mathcal{X}}(0)}$; but this would imply $\varphi(T) = 0$ for all $\varphi \in \mathcal{X}$, and then $T = 0$ by (12.30), contradicting the existence of $\psi$. So the unit balls are equal and then $\mathcal{X} = \mathcal{M}_*$.

Now we devote the rest of the proof to showing that (12.28)-(12.31) hold. The way we will use that $\mathcal{M} = \mathcal{X}^*$ is that Banach–Alaoglu applies: the closed balls in $\mathcal{M}$ are weak*-compact (Corollary 12.3.20).

*Proof of* (12.28). Suppose that $\{X_j\} \subset \mathcal{M}^{\text{sa}}$ is a net with $\|X_j\| \leq 1$ for all $j$ and $X_j \xrightarrow{\text{weak}^*} X$. We can write $X = A + iB$ with $A, B \in \mathcal{M}^{\text{sa}}$. Let $r = 1$ if $\|B\| \in \sigma(B)$ and $r = -1$ otherwise (in which case $-\|B\| \in \sigma(B)$). If $B \neq 0$, then, for $n \in \mathbb{N}$ such that $n \geq \|B\|^{-1}$,

$$\begin{aligned}
\|X_j + in\, I_{\mathcal{M}}\| &= \|(X_j - in\, I_{\mathcal{M}})(X_j + in\, I_{\mathcal{M}})\|^{1/2} = \|X_j^2 + n^2\, I_{\mathcal{M}}\|^{1/2} \\
&\leq (1 + n^2)^{1/2} = n + (1+n^2)^{1/2} - n = n + \frac{1}{(1+n^2)^{1/2} + n} \\
&< n + \frac{1}{n} \leq n + \|B\| = \|rn\, I_{\mathcal{M}} + B\| \\
&= \|\mathrm{Im}\,(A + i(B + rn\, I_{\mathcal{M}}))\| \\
&\leq \|A + i(B + rn\, I_{\mathcal{M}})\|.
\end{aligned}$$

This is a contradiction, for $X_j + in\, I_{\mathcal{M}} \xrightarrow{\text{weak}^*} A + i(B + rn\, I_{\mathcal{M}})$ and $X_j + in\, I_{\mathcal{M}} \in \overline{B^{\mathcal{M}}_{\sqrt{1+n^2}}(0)}$ which is weak*-closed, so $\|A + i(B + rn\, I_{\mathcal{M}})\| \leq (1+n^2)^{1/2}$. It follows that $B = 0$, which shows—after scaling—that $\mathcal{M}^{\text{sa}} \cap \overline{B_c^{\mathcal{M}}(0)}$ is weak*-closed for all $c > 0$. Then $\mathcal{M}^{\text{sa}}$ is weak*-closed by Krein–Šmulian.

*Proof of* (12.29) *and* (12.30). With what we just did and the fact that $T \in \overline{B_1^{\mathcal{M}}(0)}$ is positive if and only if $T^* = T$ and $0 \leq T \leq I_{\mathcal{M}}$ (Proposition 10.4.3) to get that

$$\mathcal{M}^+ \cap \overline{B_1^{\mathcal{M}}(0)} = \big(\mathcal{M}^{\text{sa}} \cap \overline{B_1^{\mathcal{M}}(0)}\big) \cap \big(I_{\mathcal{M}} - \mathcal{M}^{\text{sa}} \cap \overline{B_1^{\mathcal{M}}(0)}\big)$$

is weak*-closed, and hence weak*-compact. We also get that $\mathcal{M}^- = -\mathcal{M}^+$ is weak*-compact. Let

$$\mathcal{C} = \{Y + iZ :\ Y \in \mathcal{M}^-,\ Z \in \mathcal{M}^{\text{sa}}\}.$$

Then $\mathcal{C}$ is convex and weak*-closed. Indeed, if $\{T_j\} \subset \mathcal{C}$ and $T_j \xrightarrow{\text{weak}^*} T$, we have that $\{\mathrm{Re}\, T_j\} \subset \mathcal{M}^-$ also converges (as taking adjoints is weak*-continuous by Proposition 12.6.3). Then $\mathrm{Re}\, T = \lim \mathrm{Re}\, T_j \in \mathcal{M}^-$. So $T \in \mathcal{C}$ and $\mathcal{C}$ is weak*-closed.

If $T \in \mathcal{M}^+$ is not zero, applying Hahn–Banach (Theorem 5.7.18) and Proposition 7.2.10 to the convex weak*-compact set $\{T\}$ and the convex weak*-closed set $\mathcal{C}$, there exists $\varphi \in \mathcal{X}$ and $c \in \mathbb{R}$ with

$$\mathrm{Re}\, \varphi(S) < c < \mathrm{Re}\, \varphi(T), \qquad S \in \mathcal{C}.$$

Since $S = 0 \in \mathcal{M}^- \subset \mathcal{C}$, we get that $c > 0$ and hence $\varphi(T) \neq 0$. Also, for every $S \in \mathcal{M}^+$ we have $\operatorname{Re}\varphi(S) \geq 0$; indeed, if there existed $S_0 \in \mathcal{M}^+$ with $\operatorname{Re}\varphi(S_0) < 0$, then for $n \in \mathbb{N}$ big enough we would have $\operatorname{Re}\varphi(n(-S_0)) > c$, a contradiction since $-nS_0 \in \mathcal{C}$. For $S \in \mathcal{M}^+$ we have $inS \in \mathcal{C}$ for all $n \in \mathbb{Z}$, so $c > \operatorname{Re}\varphi(in(-S)) = \operatorname{Im}\varphi(nS)$. For $n > 0$ this gives $c/n > \operatorname{Im}\varphi(S)$ and we get $\operatorname{Im}\varphi(S) \leq 0$. Working with $n < 0$, we get $-\operatorname{Im}\varphi(S) \geq 0$. Thus $\operatorname{Im}\varphi(S) = 0$ and hence $\varphi(S) \geq 0$, so $\varphi \geq 0$. That is, for nonzero $T \in \mathcal{M}^+$ there exists $\varphi \in \mathcal{X}^+$ with $\varphi(T) > 0$.

If $T \in \mathcal{M}$ is not zero, by Exercise 11.3.9 there exist $r \in \{1, -1\}$ and $s \in \{1, i\}$ such that $(\operatorname{Re} rsT)^+ \neq 0$. This means that $rsT \notin \mathcal{C}$. Reasoning as in the previous paragraph there exists $\varphi \in \mathcal{X}^+$ with $\varphi(rsT) \neq 0$, so $\varphi(T) \neq 0$. Therefore we have shown that if $\varphi(T) = 0$ for all positive $\varphi \in \mathcal{X}$ then $T = 0$. That is, the positive functionals in $\mathcal{X}$ separate points in $\mathcal{M}$.

If $T \in \mathcal{M}$ is not positive, applying the argument from the previous paragraph we get $\varphi \in \mathcal{X}^+$ and scalars $r, s \in \{1, -1, i\}$, not both 1, such that $\varphi(rsT) > 0$; then $\varphi(T) = (rs)^{-1}\varphi(rsT)$ is not positive. Thus $T \geq 0$ if and only if $\varphi(T) \geq 0$ for all $\varphi \in \mathcal{X}^+$.

*Proof of* (12.31). Let

$$\mathcal{P} = \{\varphi_1 - \varphi_2 + i(\varphi_3 - \varphi_4) : \ \varphi_j \in \mathcal{X}^+, \ j = 1, \dots, 4\}.$$

This is a subspace of $\mathcal{X}$ (Exercise 12.7.9). Let $\psi \in \mathcal{X} \setminus \overline{\mathcal{P}}$. By Hahn–Banach (Corollary 5.7.19) and remembering that $\mathcal{X}^* = \mathcal{M}$, there exists $T \in \mathcal{M}$ with $\psi(T) = 1$ and $\varphi(T) = 0$ for all $\varphi \in \mathcal{P}$. But $\overline{\mathcal{P}}$ contains all the positive functionals in $\mathcal{X}$, so $T = 0$ by the above argument. This means that $\mathcal{X} = \overline{\mathcal{P}} = \overline{\operatorname{span}}^{\|\cdot\|} \mathcal{X}^+$. ☐

Sakai proved not only that von Neumann algebras have unique preduals, but that said property characterizes those C*-algebras which are von Neumann algebras. This requires a bit of reflection, because a C*-algebra is not necessarily represented acting on some Hilbert space. But the question can be rephrased as “which C*-algebras are isomorphic to a von Neumann algebra?” And that’s what Sakai answered. We first consider a crucial result of Kadison necessary for the proof of Sakai’s Theorem. We follow the proof from [Ped79].

**12.7.3. Proposition.** ([Kad56])

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a C*-algebra. The following statements are equivalent:*

- *(i) $\mathcal{M}$ is a von Neumann algebra;*
- *(ii) the sot-limit of every increasing bounded net in $\mathcal{M}^{\mathrm{sa}}$ is in $\mathcal{M}$.*
- *(iii) the sot-limit of every increasing bounded net of projections in $\mathcal{M}$ is in $\mathcal{M}$.*

*Proof.* (i) $\Longrightarrow$ (ii) This is Proposition 12.1.10, since by definition a von Neumann algebra is sot-closed.

(ii) $\Longrightarrow$ (iii) Trivial.

(iii) $\Longrightarrow$ (i) The proof of Corollary 12.3.10 shows that $\mathcal{M}$ is unital. This immediately shows that decreasing nets of projections also have their limit in $\mathcal{M}$. Corollary 12.4.16 implies that every $T \in \overline{\mathcal{M}}^{\text{sot}}$ is a norm limit of linear combinations of projections in $\overline{\mathcal{M}}^{\text{sot}}$. So it is enough to show that every projection $P \in \overline{\mathcal{M}}^{\text{sot}}$ is in $\mathcal{M}$. Therefore we fix a projection $P \in \overline{\mathcal{M}}^{\text{sot}}$. Let $\xi \in P\mathcal{H}$ and $\eta \in (I_{\mathcal{M}} - P)\mathcal{H}$ be unit vectors. By Kaplansky (Theorem 12.3.17) $P \in \overline{B_1^{\mathcal{M}^+}(0)}^{\text{sot}}$. So for each $n \in \mathbb{N}$ there exists $X_n \in \mathcal{M}^+$ with $\|X_n\| \leq 1$ and

$$\|\xi - X_n\xi\| < \frac{1}{n}, \qquad \|X_n\eta\| < \frac{1}{n2^n}.$$

Let $g(t) = t(1+t)^{-1}$. Via Continuous Functional Calculus (Definition 11.2.7) we define, for $m > n$,

$$Y_{n,m} = g\Big(\sum_{k=n}^{m} kX_k\Big) \in \overline{B_1^{\mathcal{M}^+}(0)}.$$

From the very coarse inequality $g(t) \leq t$ for $t \geq 0$ we get $Y_{n,m} \leq \sum_{k=n}^{m} kX_k$. Then

$$\langle Y_{n,m}\eta, \eta\rangle \leq \sum_{k=n}^{m} k\,\langle X_k\eta, \eta\rangle \leq \sum_{k=n}^{m} \frac{1}{2^k} < \frac{1}{2^{n-1}}. \tag{12.32}$$

From Exercise 11.3.7 we know that $g$ is operator monotone. As $\sum_{k=n}^{m} kX_k \geq mX_m$, we get that $Y_{n,m} \geq g(mX_m) = (I_{\mathcal{M}} + mX_m)^{-1}mX_m$. Hence

$$\begin{aligned} I_{\mathcal{M}} - Y_{n,m} &\leq I_{\mathcal{M}} - (I_{\mathcal{M}} + mX_m)^{-1}mX_m = (I_{\mathcal{M}} + mX_m)^{-1} \\ &\leq \frac{1}{m+1}\,(I_{\mathcal{M}} + m(I_{\mathcal{M}} - X_m)) \end{aligned}$$

(the last inequality is the same as $(m+1)I_{\mathcal{M}} \leq \big(I_{\mathcal{M}} + m(I_{\mathcal{M}} - X_m)\big)\big(I_{\mathcal{M}} + mX_m\big)$, which is trivial after expanding since $0 \leq X_m \leq I_{\mathcal{M}}$). Therefore

$$\begin{aligned} \langle (I_{\mathcal{M}} - Y_{n,m})\xi, \xi\rangle &\leq \frac{1}{m+1}\langle\big(I_{\mathcal{M}} + m(I_{\mathcal{M}} - X_m)\big)\xi, \xi\rangle \\ &\leq \frac{1 + m\|\xi - X_m\xi\|}{m+1} < \frac{2}{m+1}. \end{aligned} \tag{12.33}$$

For fixed $n$ the sequence $\{Y_{n,m}\}_m \subset \mathcal{M}^{\text{sa}}$ is monotone increasing and bounded (since it is in the unit ball) and so by hypothesis and Proposition 12.1.10 we have $Y_n = \lim_{\text{sot}} Y_{n,m} \in \mathcal{M}$ (the existence given by the proposition, the belonging to $\mathcal{M}$ by hypothesis). By construction $Y_{n+1,m} \leq Y_{n,m}$, so $\{Y_n\} \subset \mathcal{M}^+$ is decreasing. Again by Proposition 12.1.10 and the hypothesis, now applied to $\{I_{\mathcal{M}} - Y_n\}$, there exists $Y = \lim_{\text{sot}} Y_n \in \mathcal{M}^+$. From (12.32) and (12.33)

$$\langle Y_n\eta,\eta\rangle \le \frac{1}{2^{n-1}}, \qquad\qquad \langle (I_{\mathcal{M}}-Y_n)\xi,\xi\rangle \le 0.$$

As $0 \le Y \le I_{\mathcal{M}}$ the last inequality implies that $\langle (I_{\mathcal{M}}-Y)\xi,\xi\rangle = 0$ and the first one $\langle Y\eta,\eta\rangle = 0$. These we can write as $\|(I_{\mathcal{M}}-Y)^{1/2}\xi\| = 0$, $\|Y\eta\| = 0$. So $Y\xi = \xi$, $Y\eta = 0$.

By Corollary 12.1.16, the range projection $P_{\xi,\eta}$ of $Y$ is in $\mathcal{M}$, and we have $P_{\xi,\eta}\xi = \xi$, $P_{\xi,\eta}\eta = 0$. For $\xi_1,\xi_2 \in P\mathcal{H}$ and $\eta \in (P\mathcal{H})^\perp$, using Proposition 12.1.17 we have (since the sequence is monotone)

$$P_{\xi_1,\eta}\wedge P_{\xi_2,\eta} = \lim_{\text{sot}} (P_{\xi_1,\eta}P_{\xi_2,\eta})^n \in \mathcal{M}.$$

From this we obtain that for each $F = \{\eta_1,\ldots,\eta_n\} \subset (I_{\mathcal{M}}-P)\mathcal{H}$ the projection $Q_{F,\xi} = P_{\xi,\eta_1}\wedge\cdots\wedge P_{\xi,\eta_n}$ is in $\mathcal{M}$. If we order the finite subsets of $(I_{\mathcal{M}}-P)\mathcal{H}$ by inclusion, the net $\{Q_{F,\xi}\}$ is decreasing, and so $P_\xi = \lim_{\text{sot}} Q_{F,\xi} \in \mathcal{M}$ by hypothesis (properly, the hypothesis is applied to $\{I_{\mathcal{M}} - Q_{F,\xi}\}$). By construction, $P_\xi\xi = \xi$ and $P_\xi\eta = 0$ for all $\eta \in (P\mathcal{H})^\perp$. If we now take $F = \{\xi_1,\ldots,\xi_n\} \subset P\mathcal{H}$, let $P_F = P_{\xi_1}\vee\cdots\vee P_{\xi_n} \in \mathcal{M}$ (note that joins of projections stay in $\mathcal{M}$ by Exercise 12.1.20, since $R\vee S = I_{\mathcal{M}} - (I_{\mathcal{M}}-R)\wedge(I_{\mathcal{M}}-S)$). The net $\{P_F\}$ is increasing, and then by hypothesis $P = \lim_{\text{sot}} P_F \in \mathcal{M}$. □

*12.7.4. Remark*. In light of Proposition 12.7.3 and Proposition 12.1.10 it is tempting to think that von Neumann algebras can be abstractly characterized as those C*-algebras such that bounded monotone increasing nets of selfadjoints have a supremum. But that is not the case. Monotone complete C*-algebras exist that are not isomorphic to any von Neumann algebra. The thing is that the abstract supremum inside the algebra does not necessarily agree with the sot limit when the algebra is represented on some $\mathcal{B}(\mathcal{H})$. This became to be understood in the 1950s with attempts to characterize von Neumann algebras abstractly. Kaplansky [**Kap51a**] was the pioneer in this, with the notion of AW*-algebra (AW*, from **A**bstract W*-algebra); these are C*-algebras such that all orthogonal families of projections have a least upper bound and all maximal abelian subalgebras are monotone complete (a more abstract definition in terms of left annihilators was eventually adopted). Dixmier [**Dix51**] soon found an example of an abelian C*-algebra which is monotone complete but is not a von Neumann algebra. What happens is that if one takes $\mathcal{A} = B_b[0,1]/M[0,1]$, where $B_b[0,1]$ are the bounded Borel functions and $M[0,1]$ is the ideal of all $f \in B_b[0,1]$ such that $\{f \neq 0\}$ is meagre, this C*-algebra has no normal states. And that's the issue, an AW*-algebra has a very rich projection structure that mimics that of von Neumann algebras, but it may lack sufficiently many normal states. Another notion, monotone complete C*-algebras (where every bounded increasing net of selfadjoints admits a supremum, analog to Proposition 12.7.3) has also been extensively studied; to this day the problem of whether every AW$^*$-algebra is monotone complete is open. For a very long time since Dixmier's example it was thought that the difference between AW$^*$ and von Neumann algebras lied in the abelian part. But then Wright [Wri76] showed examples of AW$^*$-factors—so very non-abelian—which are not von Neumann algebras; these are usually called *wild factors*. Interest on monotone complete and AW$^*$-algebras waned in the 1970s after most research became focused on extending some property of von Neumann algebras to the AW$^*$-case (with Pedersen jokingly proposing a "voluntary limit of one paper per customer" in [Ped79]). Some development continued mostly through Wright–Saito and Hamana. The purely abstract theory is neatly summarized in [Ber72].

The fact that the lack of normal states is what separates AW$^*$-algebras from von Neumann algebras points, in light of Theorem 12.7.2, to the predual. And indeed the predual is the key:

**12.7.5. Theorem. (Sakai)**

*Let $\mathcal{A}$ be a C$^*$-algebra. The following statements are equivalent:*

*(i) there exists a von Neumann algebra $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ such that $\mathcal{A} \simeq \mathcal{M}$ (as C$^*$-algebras);*

*(ii) there exists a Banach space $\mathcal{X}$ such that $\mathcal{X}^* \simeq \mathcal{A}$ (isometrically as Banach spaces).*

*Proof.* (i) $\implies$ (ii) We know from Theorem 12.7.2 that $\mathcal{M} \simeq (\mathcal{M}_*)^*$ (as Banach spaces) via the duality $\hat{T}(\varphi) = \varphi(T)$. Thus we take $\mathcal{X} = \mathcal{M}_*$ and then $\mathcal{A} \simeq \mathcal{M} \simeq \mathcal{X}^*$.

(ii) $\implies$ (i) For each state $\varphi \in \mathcal{X}$ we consider its GNS representation $(\pi_\varphi, \mathcal{H}_\varphi, \xi_\varphi)$. Let

$$\mathcal{H} = \bigoplus_{\varphi \in \mathcal{X}^+,\ \|\varphi\|=1} \mathcal{H}_\varphi, \qquad \pi = \bigoplus_{\varphi \in \mathcal{X}^+,\ \|\varphi\|=1} \pi_\varphi.$$

Then $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ is a faithful representation since the states of $\mathcal{X}$ separate points (this we know from (12.30); note that we didn't use that $\mathcal{M}$ was a von Neumann algebra in the proof, only that it has a predual). We want to show that $\pi(\mathcal{A})$ is a von Neumann algebra. Let $\{a_j\} \subset \mathcal{A}^{\mathrm{sa}}$ be a bounded increasing net, and let $T = \lim_{\mathrm{sot}} \pi(a_j) \in \mathcal{B}(\mathcal{H})$. By (12.28) we have that $\mathcal{A}^{\mathrm{sa}} \cap \overline{B_1^{\mathcal{A}}(0)}$ is weak$^*$-compact, so $\{a_j\}$ admits a weak$^*$-convergent subnet $\{a_{j_k}\}$ with $a_{j_k} \xrightarrow{\text{weak}^*} a$. Given any index $j$ there will be $k$ such that $j_k \geq j$.

So

$$\varphi(a - a_j) = \lim_k \varphi(a_{j_k} - a_j) \geq 0$$

for all $\varphi \in \mathcal{X}^+$. Then by (12.29) we have that $a \geq a_j$, so $a$ is an upper bound for $\{a_j\}$. If $b \in \mathcal{A}^{\mathrm{sa}}$ is another upper bound for $\{a_j\}$, then $\varphi(b-a) = \lim_k \varphi(b - a_{j_k}) \geq 0$ for all $\varphi \in \mathcal{X}^+$, and thus $b \geq a$. Hence $a$ is the least upper bound for $\{a_j\}$. As mentioned in Remark 12.7.4, the fact that $a = \sup a_j$ and $T = \sup\{\pi(a_j) : j\}$ does not guarantee in itself that $\pi(a) = T$. But we have a particular representation here. Indeed, for any state $\varphi \in \mathcal{X}^+$ we have

$$\langle \pi_\varphi(a_j)\pi_\varphi(b)\xi_\varphi, \pi_\varphi(c)\xi_\varphi\rangle = \varphi(c^*a_jb) \to \varphi(c^*ab) = \langle \pi_\varphi(a)\pi_\varphi(b)\xi_\varphi, \pi_\varphi(c)\xi_\varphi\rangle,$$

and then

$$\langle (T - \pi(a))\pi_\varphi(b)\xi_\varphi, \pi_\varphi(c)\xi_\varphi\rangle = \lim_j \langle (T - \pi(a_j))\pi_\varphi\xi_\varphi, \pi_\varphi(c)\xi_\varphi\rangle = 0$$

by definition of $T$. Given $\varepsilon > 0$ and $\eta \in \mathcal{H}$, from $\sum_\varphi \|\eta_\varphi\|^2 = \|\eta\|^2 < \infty$ we get that there exist countably many $\{\varphi_n\}$ such that $\|\eta_{\varphi_n}\| \neq 0$. For each $n$ we can choose $b_n \in \mathcal{A}$ such that $\|\eta_{\varphi_n} - \pi(b_n)\xi_\varphi\| < 2^{-n}\varepsilon$. Then

$$\begin{aligned}
\langle (T - \pi(a))\eta, \eta\rangle &= \sum_n \langle (T - \pi(a))\eta_{\varphi_n}, \eta_{\varphi_n}\rangle \\
&\leq \big(\|T\| + \|\pi(a)\|\big)\,\big(2\|\eta\| + \varepsilon\big)\,\varepsilon \\
&\qquad + \sum_\varphi \langle (T - \pi(a))\pi(b_n)\xi_\varphi, \pi(b_n)\xi_\varphi\rangle \\
&= \big(\|T\| + \|\pi(a)\|\big)\,\big(2\|\eta\| + \varepsilon\big)\,\varepsilon.
\end{aligned}$$

As this happens for all $\varepsilon > 0$ and $\eta \in \mathcal{H}$, via polarization we get $\pi(a) = T$. Then Proposition 12.7.3 shows that $\pi(\mathcal{A})$ is a von Neumann algebra. □

The predual of a C*-algebra is necessarily unique (Exercise 12.7.10).

**12.7.2. The word "separable".** While for Banach spaces and Banach algebras and C*-algebras the word "separable" has a clear meaning (a countable dense set exists for the obvious topology) the word is more loaded when talking about von Neumann algebras. It is common in the literature to require a von Neumann algebra to

- have separable predual; or
- act on a separable Hilbert space; or
- be separable.

It is easy to be confused here because the qualifier *separable* is used in the first two cases with respect to the norm (of the predual and of the Hilbert space, respectively), while in the third case it is used with reference to the sot/wot topologies. These assertions are related in the the following way: for $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$,

$$\mathcal{H} \text{ separable} \overset{(\Longleftarrow)}{\Longrightarrow} \mathcal{M} \text{ has separable predual} \implies \mathcal{M} \text{ separable.} \tag{12.34}$$

The converse of the first implication is true in the sense that if $\mathcal{M}_*$ is separable, then there is some $\sigma$-weakly continuous faithful representation $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{K})$ with $\mathcal{K}$ separable. To visualize what could happen take any non-separable $\mathcal{H}$ and let $\mathcal{M} = \mathbb{C}\, I_{\mathcal{H}}$. Then $\mathcal{M}$ is as small as can be, and it admits a faithful normal representation onto a one-dimensional $\mathcal{K}$.

It is not unusual to see the phrase "$\mathcal{M}$ separable" to refer to the first two properties and not to the third one, at least informally. But the most common assertion in the classical von Neumann algebra literature is to request a separable predual.

Let us now check the relations (12.34). If $\mathcal{H}$ is separable, using Theorem 12.7.2 to say that the predual is unique, via Corollary 7.3.8 we have $\mathcal{M}_* \simeq \mathcal{T}(\mathcal{H})/\mathcal{M}_o$, which is separable ($\mathcal{T}(\mathcal{H})$ is separable and the quotient map is contractive, so a dense subset will stay dense in the quotient).

Conversely, if $\mathcal{M}_*$ is separable, let $\{\varphi_n\} \subset \mathcal{M}_*$ a countable dense subset. By Proposition 12.6.3 each $\varphi_n$ is defined in terms of two countable sequences $\{\xi_{n,m}\}, \{\eta_{n,m}\} \subset \mathcal{H}$. Then $\mathcal{K} = \overline{\operatorname{span}}^{\|\cdot\|}\{\xi_{n,m}, \eta_{n,m} :\ n, m \in \mathbb{N}\}$ is separable. If $T \in \mathcal{M}$ and $T|_{\mathcal{K}} = 0$, it follows that $\varphi_n(T) = 0$ for all $n$ and thus $\varphi(T) = 0$ for all $\varphi \in \mathcal{M}_*$, which means that $T = 0$. So the representation $T \longmapsto T|_{\mathcal{K}}$ is normal and faithful, and it maps $\mathcal{M}$ to a von Neumann algebra acting on a separable Hilbert space.

And when $\mathcal{M}_*$ is separable, by the previous paragraph we can represent $\mathcal{M}$ acting on a separable space; so we may assume that $\mathcal{H}$ is separable. The sot is metrizable on the unit ball by Exercise 12.1.3, so the unit ball of $\mathcal{M}$ is separable as it is a subset of a separable topological space (Proposition 1.8.5). Hence $\mathcal{M}$ is separable.

Finally, von Neumann algebras exist which are separable but have non-separable predual. One such example is $\mathcal{M} = \ell^\infty(\mathbb{R})$ acting on $\ell^2(\mathbb{R})$. Its predual is $\ell^1(\mathbb{R})$ (by Exercise 5.6.11). This is not separable, as it contains the uncountable set $\{e_t\}$ (the canonical basis) with $\|e_t - e_s\|_1 = 1$ for all $s \neq t$. It remains to see that $\ell^\infty(\mathbb{R})$ is $\sigma$-weak separable. The closed unit ball of $\mathcal{M}$ consists of the set $A$ of all the functions $f : \mathbb{R} \to \overline{\mathbb{D}}$. On $A$, the $\sigma$-weak topology agrees with pointwise convergence (Exercise 12.7.7), and this allows one to show that $\ell^\infty(\mathbb{R})$ is $\sigma$-weak separable (Exercise 12.7.8).

### 12.7.3. Universal Representation and Enveloping von Neumann Algebra.

**12.7.6. Definition.**

*Let $\mathcal{A}$ be a $C^*$-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. We say that $\pi$ is* ***universal*** *if any representation $\rho : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ extends to a $\sigma$-weak continuous $*$-homomorphism $\tilde{\rho} : \pi(\mathcal{A})'' \to \rho(\mathcal{A})''$ such that $\tilde{\rho} \circ \pi = \rho$.*

*When $\pi$ is universal, the von Neumann algebra $\pi(\mathcal{A})''$ is called the* ***enveloping von Neumann algebra*** *of $\mathcal{A}$.*

The enveloping von Neumann algebra of $\mathcal{A}$ is well-defined, in that it is uniquely determined up to isomorphism. Concretely, if $\pi_1 : \mathcal{A} \to \mathcal{B}(\mathcal{H}_1)$ and $\pi_2 : \mathcal{A} \to \mathcal{B}(\mathcal{H}_2)$ are universal, there exist $\tilde{\pi}_2 : \pi_1(\mathcal{A})'' \to \pi_2(\mathcal{A})''$ and $\tilde{\pi}_1 : \pi_2(\mathcal{A})'' \to \pi_1(\mathcal{A})''$ with $\tilde{\pi}_2 \circ \pi_1 = \pi_2$, and $\tilde{\pi}_1 \circ \pi_2 = \pi_1$. Then

$$\tilde{\pi}_2 \circ \tilde{\pi}_1 \circ \pi_2 = \tilde{\pi}_2 \circ \pi_1 = \pi_2,$$

and similarly $\tilde{\pi}_1 \circ \tilde{\pi}_2 \circ \pi_1 = \pi_1$. That is, $\tilde{\pi}_2 \circ \tilde{\pi}_1$ is the identity on $\pi_2(\mathcal{A})$ and $\tilde{\pi}_1 \circ \tilde{\pi}_2$ is the identity on $\pi_1(\mathcal{A})$. As $\tilde{\pi}_1, \tilde{\pi}_2$ are $\sigma$-weakly continuous and $\pi_1(\mathcal{A})$ and $\pi_2(\mathcal{A})$ are respectively $\sigma$-weakly dense in $\pi_1(\mathcal{A})''$ and $\pi_2(\mathcal{A})''$, via Corollary 12.6.14 we get that $\pi_1(\mathcal{A})''$ and $\pi_2(\mathcal{A})''$ are isomorphic.

It is not automatically obvious—though true, as we will see—that universal representations exist.

***12.7.7. Proposition.***

*Let $\mathcal{A}$ be a $C^*$-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. Then there exists a unique linear map $\tilde{\pi} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ such that*

*(i) $\tilde{\pi} \circ \iota = \pi$, where $\iota$ is the canonical inclusion $\iota : \mathcal{A} \to \mathcal{A}^{**}$;*

*(ii) $\tilde{\pi}$ is weak$^*$ to $\sigma$-weak continuous;*

*(iii) $\tilde{\pi}\big(\overline{B_1^{\mathcal{A}^{**}}(0)}\big) = \overline{B_1^{\pi(\mathcal{A})''}(0)}$.*

*Proof.* By Theorem 12.7.2, $\pi(\mathcal{A})''$ is the dual of the space $\pi(\mathcal{A})''_*$ of its normal functionals. Let $\pi_0 : \pi(\mathcal{A})''_* \to \mathcal{A}^*$ be the restriction of the adjoint $\pi^*$ to $\pi(\mathcal{A})''_*$, and define $\tilde{\pi} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ by $\tilde{\pi} = \pi_0^*$. Then, for any $a \in \mathcal{A}$ and $\varphi \in \pi(\mathcal{A})''_*$,

$$\begin{aligned}(\tilde{\pi}(\iota(a)))(\varphi) &= (\pi_0^*(\iota(a)))(\varphi) = (\iota(a))(\pi^*(\varphi)) \\ &= (\pi^*(\varphi))(a) = \varphi(\pi(a)) = (\pi(a))(\varphi).\end{aligned}$$

Hence $\tilde{\pi} \circ \iota = \pi$. If $\Psi_j \xrightarrow{\text{weak}^*} 0$ in $\mathcal{A}^{**}$ and $\phi \in \pi(\mathcal{A})''_*$, then

$$(\tilde{\pi}(\Psi_j))\phi = (\pi_0^*(\Psi_j))\phi = (\Psi_j \circ \pi^*)\phi = \Psi_j(\pi^* \circ \phi) \to 0,$$

so $\tilde{\pi}$ is weak$^*$ to $\sigma$-weak continuous. Since $\pi$ is a contraction (Proposition 11.4.9) it maps $B_1^{\mathcal{A}}(0)$ into $B_1^{\pi(\mathcal{A})}(0)$, The $*$-isomorphism $\mathcal{A}/\ker\pi \to$

$\pi(\mathcal{A})$ induced by $\pi$ is isometric (again by Proposition 11.4.9). Fix $a \in \mathcal{A}$ with $\|\pi(a)\| \leq 1$. By Proposition 11.7.6 there exists $b \in \mathcal{A}$ with $b + \ker\pi = a + \ker\pi$ and $\|b\| = \|a + \ker\pi\| = \|\pi(a)\| \leq 1$. Hence $\pi(a) = \pi(b)$ with $\|b\| \leq 1$, showing that $\pi : \overline{B_1^{\mathcal{A}}(0)} \to \overline{B_1^{\pi(\mathcal{A})}(0)}$ is onto. Now we can show (iii):

$$\tilde{\pi}\big(\overline{B_1^{\mathcal{A}^{**}}(0)}\big) \underset{\text{(Theorem 7.2.18)}}{=} \tilde{\pi}\Big(\overline{B_1^{\iota(\mathcal{A})}(0)}^{w^*}\Big) \overset{(\iota \text{ is isometric})}{=} \tilde{\pi}\Big(\overline{\iota\big(B_1^{\mathcal{A}}(0)\big)}^{w^*}\Big) \underset{\text{(Alaoglu and Exercise 1.8.40)}}{=} \overline{\tilde{\pi}\circ\iota\big(B_1^{\mathcal{A}}(0)\big)}^{w^*}$$

$$= \overline{\pi\big(B_1^{\mathcal{A}}(0)\big)}^{w^*} = \overline{B_1^{\pi(\mathcal{A})}(0)}^{w^*} \underset{\text{(Lemma 12.1.21)}}{=} \overline{B_1^{\pi(\mathcal{A})}(0)}^{\text{wot}} \overset{\text{(Corollary 12.1.3)}}{=} \overline{B_1^{\pi(\mathcal{A})}(0)}^{\text{sot}}$$

$$\underset{\text{(Kaplansky)}}{=} \overline{B_1^{\overline{\pi(\mathcal{A})}^{\text{sot}}}(0)} \underset{\text{(Double Commutant Theorem)}}{=} \overline{B_1^{\pi(\mathcal{A})''}(0)}. \qquad \square$$

***12.7.8. Theorem. (Universal Representation and Enveloping von Neumann Algebra)***

*Let $\mathcal{A}$ be a $C^*$-algebra. There exists a universal representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$. For such a representation there is a unique linear surjective isometry $\tilde{\pi} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ which satisfies $\tilde{\pi}\circ\iota = \pi$, and it is a weak$^*$–$\sigma$-weak homeomorphism.*

*Proof.* We take $\pi$ to be the faithful representation that was constructed in Corollary 11.6.6, and let $\tilde{\pi} : \mathcal{A}^{**} \to \pi(\mathcal{A})''$ as in Proposition 12.7.7. Given $\varphi \in \mathcal{A}^*$ we can write $\varphi = \langle \pi_\psi(\cdot)\xi, \eta\rangle$ by Proposition 11.6.9 for some $\psi \in S(\mathcal{A})$ and $\xi, \eta \in \mathcal{H}_\psi$, and by Remark 11.6.4 we may assume that $\psi \in \mathcal{F}$ in the sense of the proof of Corollary 11.6.6. If we form $\tilde{\xi}, \tilde{\eta} \in \bigoplus_{\phi\in\mathcal{F}} \mathcal{H}_\phi$ as $\tilde{\xi} = \delta_\psi\,\xi$, $\tilde{\eta} = \delta_\psi\,\eta$, then

$$\langle \pi(a)\tilde{\xi}, \tilde{\eta}\rangle = \langle \pi_\psi(a)\xi, \eta\rangle = \varphi(a).$$

So in the notation of the proof of Proposition 12.7.7, if $\phi$ is the normal linear functional induced by $\pi(a) \longmapsto \langle \pi(a)\tilde{\xi}, \tilde{\eta}\rangle$ we have that $\phi \in \pi(\mathcal{A})''_*$ and (with $\pi_0$ as in the proof of Proposition 12.7.7)

$$(\pi_0(\phi))(a) = \phi(\pi(a)) = \varphi(a).$$

This shows that $\pi_0$ is onto, and by Corollary 9.4.9 we get that $\tilde{\pi}$ is injective. It is surjective by (iii) in Proposition 12.7.7, and isometric by said property and Exercise 9.1.5. Also (iii) of Proposition 12.7.7 guarantees that $\tilde{\pi}$ is bicontinuous as it sends closed balls to closed balls.

It remains to check the universality. Given a representation $\rho : \mathcal{A} \to \mathcal{B}(\mathcal{K})$, let $\tilde{\rho} : \mathcal{A}^{**} \to \rho(\mathcal{A})''$ be the corresponding map from Proposition 12.7.7. Put $\pi_\rho : \pi(\mathcal{A})'' \to \rho(\mathcal{A})''$ to be $\pi_\rho = \tilde{\rho} \circ \tilde{\pi}^{-1}$. Then $\pi_\rho$ is $\sigma$-weak continuous and

$$\pi_\rho \circ \pi(a) = \tilde{\rho}(\tilde{\pi}^{-1}(\pi(a))) = \tilde{\rho}(\iota(a)) = \rho(a).$$

The equality also shows that $\pi_\rho$ is multiplicative and $*$-preserving on $\pi(\mathcal{A})$, and hence on $\pi(\mathcal{A})''$ by the $\sigma$-weak continuity. Thus $\pi$ is universal. $\square$

So now we know that universal representations exist for any C$^*$-algebra $\mathcal{A}$, and so does the enveloping von Neumann algebra. And the enveloping von Neumann algebra of $\mathcal{A}$ has a canonical identification with $\mathcal{A}^{**}$, and the correspondence is linear, isometric, and homeomorphic if we consider the $\sigma$-weak topology in the enveloping von Neumann algebra and the weak$^*$-topology in $\mathcal{A}^{**}$. Because this latter space does not depend on a concrete representation, it is common in the literature to say "the double dual $\mathcal{A}^{**}$ of $\mathcal{A}$" to refer to the enveloping von Neumann algebra of $\mathcal{A}$; some authors even use $\mathcal{A}''$. We will also use this terminology occasionally.

The universal representation is very non-unique, as the direct sum of representations used in Corollary 11.6.6 can be constructed out of any family of states that separates points. But the uniqueness in Theorem 12.7.8 guarantees that all are equivalent. In particular we can see, in light of Corollary 11.6.6, that any linear functional $\varphi \in \mathcal{A}^*$ extends to a normal linear functional on $\mathcal{A}^{**}$ (we can see this either from the fact that $\varphi$ is weak$^*$-continuous on $\mathcal{A}^{**}$, or from the fact that it is of the form $\varphi(a) = \langle a\xi, \eta\rangle$).

There is a way to define the multiplicative structure in $\mathcal{A}^{**}$ without recurring to the above results on representations. One can construct what is called the **Arens Product** on $\mathcal{A}^{**}$.

**12.7.4. Exercises.**

(12.7.1) Let $\mathcal{M}$ be a von Neumann algebra. Show that $\mathcal{M}_*$ is norm-closed.

(12.7.2) Provide an alternative proof to the fact that $\mathcal{M}_*$ is a predual for the von Neumann algebra $\mathcal{M}$ by using Corollary 7.3.8 to see that a predual for $\mathcal{M}$ is given by $\mathcal{T}(\mathcal{H})/\mathcal{M}_o$. This means identifying all normal functionals that agree on $\mathcal{M}$, so we have precisely the normal functionals of $\mathcal{M}$.

(12.7.3) Let $\mathcal{A}, \mathcal{B}$ be C$^*$-algebras and $\gamma : \mathcal{A} \to \mathcal{B}$ a $*$-isomorphism. Show that $\mathcal{A}^{**} \simeq \mathcal{B}^{**}$ as von Neumann algebras.

(12.7.4) Let $\mathcal{A}$ be a unital C$^*$-algebra. Show that the unit of $\mathcal{A}^{**}$ is the unit of $\mathcal{A}$; that is, show that $\hat{I}_\mathcal{A} = I_{\mathcal{A}^{**}}$.

(12.7.5) Let $\mathcal{A}$ be a C$^*$-algebra and $Z \in \mathcal{Z}(\mathcal{A}^{**})$ a central projection. Show that $(Z\mathcal{A})^{**} = Z\mathcal{A}^{**}$.

(12.7.6) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{J} \subset \mathcal{A}$ a proper ideal. Show that $\mathcal{A}^{**} \simeq \mathcal{J}^{**} \oplus (\mathcal{A}/\mathcal{J})^{**}$ as C*-algebras.

(12.7.7) Let $A$ be the closed unit ball of $\ell^\infty(\mathbb{R}) \subset \mathcal{B}(\ell^2(\mathbb{R}))$. Show that on $A$ the $\sigma$-weak topology is precisely pointwise convergence.

(12.7.8) Show that $\ell^\infty(\mathbb{R}) \subset \mathcal{B}(\ell^2(\mathbb{R}))$ is $\sigma$-weak separable.

(12.7.9) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{X}$ a Banach space such that $\mathcal{M} = \mathcal{X}^*$. Put $\mathcal{P} = \{\varphi_1 - \varphi_2 + i(\varphi_3 - \varphi_4) : \ \varphi_j \in \mathcal{X}^+, \ j = 1, \ldots, 4\}$. Show that $\mathcal{P}$ is a subspace.

(12.7.10) Let $\mathcal{A}$ be a C*-algebra and $\mathcal{X}, \mathcal{Y}$ Banach spaces such that $\mathcal{X}^* = \mathcal{Y}^* = \mathcal{A}$. Show that $\mathcal{X} \simeq \mathcal{Y}$ isometrically.

CHAPTER 13

# Constructions with C*-Algebras

Before talking more about von Neumann algebras, we will focus on some general techniques and constructions. For this, after considering algebraic tensor products we will venture into discussing completely positive maps before going on with group C*-algebras and crossed products, and eventually topological tensor products.

## 13.1. Algebraic Tensor Products

We have already seen ways to construct new objects by somehow putting together existing objects, with direct sums the most used such construction. Tensor products appear as another natural way to aggregate objects. We will initially construct tensor products of vector spaces, but it is always worth keep in mind that Hilbert spaces, Banach spaces, Banach algebras, C*-algebras, and von Neumann algebras are all vector spaces. Eventually we will add more algebraic and topological structure to the vector space tensor product.

The idea for the tensor product is that if we have $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ we want to form the elementary tensor $x \otimes y$, that should behave as a product (it should distribute with respect to addition, for instance). The construction cannot possibly mix $x$ and $y$, since they are likely very different objects; to mention an example that will actually appear in the text, it is possible that $x$ is a matrix and $y$ is a continuous function.

**13.1.1. Definition.**

*Given vector spaces $\mathcal{X}$, $\mathcal{Y}$, we denote their* ***tensor product*** *by $\mathcal{X} \otimes \mathcal{Y}$. This is defined as*

$$\mathcal{X} \otimes \mathcal{Y} = \operatorname{span}\{x \otimes y : \; x \in \mathcal{X}, \; y \in \mathcal{Y}\}.$$

*The elements of the form $x \otimes y$ are called the* ***elementary tensors*** *and they are defined as linear maps $x \otimes y : BL(\mathcal{X}, \mathcal{Y}) \to \mathbb{C}$, where $BL(\mathcal{X}, \mathcal{Y}) = \{\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}, \quad \text{bilinear}\}$, by*

$$x \otimes y\,(\phi) := \phi(x, y).$$

*That is, we see $\mathcal{X} \otimes \mathcal{Y}$ as a subspace of the algebraic dual $BL(\mathcal{X}, \mathcal{Y})^*$.*

The definition above is not the most intuitive, but it has the advantage that it constructs $\mathcal{X} \otimes \mathcal{Y}$ explicitly. For another formal definition we suggest to look at [BO08, Section 3.1]. For other examples, the notion of tensor product of Hilbert spaces is carefully developed in [KR97a, Theorem 2.6.4].

***13.1.2. Proposition.***

*The elementary tensors satisfy the following relations: for $x, z \in X$, $y, w \in Y$, $\lambda \in \mathbb{C}$,*

*(i)* $\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y)$,

*(ii)* $(x + z) \otimes y = x \otimes y + z \otimes y$

*(iii)* $x \otimes (y + w) = x \otimes y + x \otimes w$.

*Proof.* Exercise 13.1.1. □

It is very important to keep in mind that $\mathcal{X} \otimes \mathcal{Y}$ always contains way more than the elementary tensors; it contains all linear combinations of them. Often times properties can be proven on the elementary tensors—if the property is linear—but many other times elementary tensors are not enough; this often makes results about tensor products non-trivial.

***13.1.3. Proposition.***

*If $x_1, \ldots, x_n \in \mathcal{X}$ with $x_j \neq 0$ for all $j$, and $y_1, \ldots, y_n \in \mathcal{Y}$ are linearly independent, then $x_1 \otimes y_1, \ldots, x_n \otimes y_n$ are linearly independent.*

*Proof.* Suppose that $\sum_j \alpha_j\, x_j \otimes y_j = 0$. Complete $\{y_j\}$ to a basis of $\mathcal{Y}$, and let $g_k(y)$ be the linear map that sends $y$ to the coefficient of $y$ corresponding to $y_k$, i.e. $y = \sum g_k(y)\, y_k$. Fix $f : \mathcal{X} \to \mathbb{C}$ linear, and define $\phi_k : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}$ by $\phi_k(x, y) = f(x) g_k(y)$. Then

$$0 = \Big(\sum_j \alpha_j x_j \otimes y_j\Big)(\phi_k) = \sum_j \alpha_j \phi_k(x_j, y_j) = \sum_j \alpha_j f(x_j) g_k(y_j) = \alpha_k f(x_k).$$

By choosing $f$ such that $f(x_k) = 1$, we deduce that $\alpha_k = 0$. We can do this for all $k$, so $\alpha_1 = \cdots = \alpha_n = 0$. □

Proposition 13.1.3 is useful in itself, as we will see, but it also helps us with addressing a basic problem with tensor products, which is that the same element of $\mathcal{X} \otimes \mathcal{Y}$ can be expressed in many ways. For a very basic example of this,

$$x \otimes y = (x/2) \otimes y + x \otimes (y/2) = (x + 4z) \otimes (y/5) + (x - z) \otimes (4y/5).$$

This means that defining functions on the tensor product requires us to check for well-definedness. This will be usually addressed via the following criterion:

**13.1.4. Proposition.**

*Let $x_1, \ldots, x_n \in \mathcal{X}$, $y_1, \ldots, y_n \in \mathcal{Y}$. The following statements are equivalent:*

(i) $0 = \displaystyle\sum_{j=1}^n x_j \otimes y_j$;

(ii) *there exists an $n \times n$ complex matrix $C$ such that $XC = 0$ and $CY = Y$, where $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)^\top$. That is,*
$$\sum_{k=1}^n c_{kj} x_k = 0 \quad \text{for all } j, \qquad \sum_{j=1}^n c_{kj} y_j = y_k \quad \text{for all } k.$$

*Proof.* (i) $\Longrightarrow$ (ii) Let $z_1, \ldots, z_m$ be a basis of $\operatorname{span}\{y_1, \ldots, y_n\}$. Then there exist coefficients such that

$$y_j = \sum_{k=1}^m \alpha_{jk} z_k, \quad z_k = \sum_{j=1}^n \beta_{kj} y_j,$$

Write $A = [\alpha_{jk}]$, $B = [\beta_{kj}]$, and define $C = AB$. Then

$$0 = \sum_{j=1}^n x_j \otimes y_j = \sum_{j=1}^n x_j \otimes \Big(\sum_{k=1}^m \alpha_{jk} z_k\Big) = \sum_{k=1}^m \Big(\sum_{j=1}^n \alpha_{jk} x_j\Big) \otimes z_k.$$

By Proposition 13.1.3, $\sum_{j=1}^{n} \alpha_{jk} x_j = 0$ for each $k$. So $XA = 0$, and then $XC = XAB = 0$. Also, for each $k$,

$$(CY)_k = \sum_{j=1}^{n} c_{kj} y_j = \sum_{j=1}^{n} \sum_{h=1}^{m} \alpha_{kh} \beta_{hj} y_j = \sum_{h=1}^{m} \alpha_{kh} \sum_{j=1}^{n} \beta_{hj} y_j = \sum_{h=1}^{m} \alpha_{kh} z_h = y_k,$$

so $CY = Y$.

(ii) $\implies$ (i) If the matrix $C$ exists, then

$$\begin{aligned}\sum_{j=1}^{n} x_j \otimes y_j &= \sum_{j=1}^{n} x_j \otimes \Big( \sum_{h=1}^{n} c_{jh} y_h \Big) = \sum_{h=1}^{n} \Big( \sum_{j=1}^{n} x_j c_{jh} \Big) \otimes y_h \\ &= \sum_{h=1}^{n} (XC)_h \otimes y_h = 0.\end{aligned}$$

□

*13.1.5. Remark .* There is a useful idea in the proof of Proposition 13.1.4: any element of $\mathcal{X} \otimes \mathcal{Y}$ can be written as $\sum_{j=1}^{m} x_j \otimes y_j$ (obviously) but we may assume, without loss of generality, that $y_1, \ldots, y_m$ are linearly independent. Or, alternatively, we may assume that $x_1, \ldots, x_m$ are. And, once any of those two choices is made, the presentation of an element as $\sum_{j=1}^{n} x_j \otimes y_j$ is unique.

The algebraic tensor product was constructed in such a way that it turns bilinear maps $\mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ into linear maps $\mathcal{X} \otimes \mathcal{Y} \to \mathcal{Z}$. This is a characterizing universal property of the tensor product.

***13.1.6. Theorem. (Universal Property of the Tensor Product)***

*Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be bilinear. Then there exists $T_\phi : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{Z}$ linear, such that $T_\phi(x \otimes y) = \phi(x, y)$ for all $x \in \mathcal{X}$, $y \in \mathcal{Y}$. Conversely, if $L : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{Z}$ is linear, there exists $\psi_L : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, bilinear, such that $\psi_L(x, y) = L(x \otimes y)$. We have, for every bilinear $\phi : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$, $\psi_{T_\phi} = \phi$; and, for every linear $L : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{Z}$, $T_{\psi_L} = L$.*

*Proof.* We want to define

$$T_\phi\Big( \sum_j x_j \otimes y_j \Big) = \sum_j \phi(x_j, y_j).$$

The only issue is whether this is well-defined. For this, if $\sum_j x_j \otimes y_j = 0$, then by Proposition 13.1.4 there exist numbers $\{c_{kj}\}$ such that

$$\sum_k c_{kj} x_k = 0, \quad \sum_j c_{kj} y_j = y_k.$$

Then

$$\sum_k \phi(x_k, y_k) = \sum_k \phi(x_k, \sum_j c_{kj} y_j) = \sum_j \phi(\sum_k c_{kj} x_k, y_j) = \sum_j \phi(0, y_j) = 0.$$

We can define $\psi_L$ by $\psi_L(x, y) = L(x \otimes y)$ and it is clearly bilinear. Finally,

$$\psi_{T_\phi}(x, y) = T_\phi(x \otimes y) = \phi(x, y),$$

and

$$T_{\psi_L}\Big(\sum_j x_j \otimes y_j\Big) = \sum_j \psi_L(x_j, y_j) = \sum_j L(x_j \otimes y_j) = L\Big(\sum_j x_j \otimes y_j\Big). \qquad \square$$

One application of Theorem 13.1.6 is that the different constructions of $\mathcal{X} \otimes \mathcal{Y}$ mentioned above produce canonically isomorphic vector spaces (Exercise 13.1.3). Another one is to show that $\mathcal{X} \otimes \mathcal{Y}$ is canonically isomorphic to $\mathcal{Y} \otimes \mathcal{X}$ (Exercise 13.1.4), and $\mathbb{C} \otimes \mathcal{X} \simeq \mathcal{X}$, $M_n(\mathbb{C}) \otimes \mathcal{X} \simeq M_n(\mathcal{X})$, also canonically (Exercise 13.1.5). Yet another application is that tensor products are associative:

***13.1.7. Proposition.***

*Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be vector spaces. Then the assignment $x \otimes (y \otimes z) \longmapsto (x \otimes y) \otimes z$ induces a canonical isomorphism between $(\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$ and $\mathcal{X} \otimes (\mathcal{Y} \otimes \mathcal{Z})$. When $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are algebras, the isomorphism induced is also an algebra isomorphism.*

*Proof.* Fix $x \in \mathcal{X}$, and consider the bilinear map $\phi^x : \mathcal{Y} \times \mathcal{Z} \to (\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$ given by $\phi^x(y, z) = (x \otimes y) \otimes z$. By Theorem 13.1.6, $T_{\phi^x} : \mathcal{Y} \otimes \mathcal{Z} \to (\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$ is a linear map such that $T_{\phi^x}(y \otimes z) = (x \otimes y) \otimes z$.

Now define a bilinear map $\phi : \mathcal{X} \times (\mathcal{Y} \otimes \mathcal{Z}) \to (\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$ by $\phi(x, w) = T_{\phi^x}(w)$. Then, again by Theorem 13.1.6, we get a linear map $T_\phi : \mathcal{X} \otimes (\mathcal{Y} \otimes \mathcal{Z}) \to (\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$. Given $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, we have

$$T_\phi(x \otimes (y \otimes z)) = \phi(x, y \otimes z) = T_{\phi^x}(y \otimes z) = \phi^x(y, z) = (x \otimes y) \otimes z.$$

Via an analog construction, we can get $T_\psi : (\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z} \to \mathcal{X} \otimes (\mathcal{Y} \otimes \mathcal{Z})$, linear, that maps $(x \otimes y) \otimes z \longmapsto x \otimes (y \otimes z)$, and so it is an inverse for $T_\phi$.

When the vector spaces involved are also algebras, the product in the tensor product is defined on elementary tensors by $(x_1 \otimes y_1)(x_2 \otimes y_2) = x_1 x_2 \otimes y_1 y_2$. As $T_\phi$ operates on elementary tensors by "moving the brackets" and it is linear, it is clear that it is multiplicative. $\square$

Since the tensor product is associative by Proposition 13.1.7 we can talk about $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$ by thinking about it as $(\mathcal{X} \otimes \mathcal{Y}) \otimes \mathcal{Z}$. We can also similarly consider the tensor product of any finite family of vector spaces.

Let us state more consequences of the universal property.

***13.1.8. Corollary. (Tensor Product of Linear Maps)***

*If $\varphi : \mathcal{X} \to \mathcal{W}$ and $\psi : \mathcal{Y} \to \mathcal{Z}$ are linear, then there exists a unique linear map $\varphi \otimes \psi : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{W} \otimes \mathcal{Z}$ such that $(\varphi \otimes \psi)(x \otimes y) = \varphi(x) \otimes \psi(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

*When $\mathcal{X}, \mathcal{Y}, \mathcal{W}, \mathcal{Z}$ are algebras and $\varphi, \psi$ are homomorphisms, $\varphi \otimes \psi$ is a homomorphism.*

*Proof.* Via Theorem 13.1.6 we define $\varphi \otimes \psi = T_b$, where $b : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{W} \otimes \mathcal{Z}$ is the bilinear form $b(x, y) = \varphi(x) \otimes \psi(y)$. If $L : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{W} \otimes \mathcal{Z}$ is linear and $L(x \otimes y) = \varphi(x) \otimes \psi(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, the bilinear form induced by $L$ is $b_L(x, y) = L(x \otimes y) = \varphi(x) \otimes \psi(y)$; it agrees with $b$, and so $L = T_{b_L} = T_b = \varphi \otimes \psi$.

When we have algebras and homomorphisms, because $\varphi \otimes \psi$ is linear it is enough to check for multiplicativity on elementary tensors. We have

$$\begin{aligned}
(\varphi \otimes \psi)\big((x_1 \otimes y_1)(x_2 \otimes y_2)\big) &= (\varphi \otimes \psi)(x_1 x_2 \otimes y_1 y_2) = \varphi(x_1 x_2) \otimes \psi(y_1 y_2) \\
&= \varphi(x_1)\varphi(x_2) \otimes \psi(y_1)\psi(y_2) \\
&= (\varphi(x_1) \otimes \psi(y_1))(\varphi(x_2) \otimes \psi(y_2)) \\
&= (\varphi \otimes \psi)(x_1 \otimes y_1)\,(\varphi \otimes \psi)(x_2 \otimes y_2).
\end{aligned}$$

□

***13.1.9. Corollary.***

*Let $\mathcal{X}, \mathcal{Y}$ be vector spaces and $\mathcal{A}$ an algebra. If $\varphi : \mathcal{X} \to \mathcal{A}$ and $\psi : \mathcal{Y} \to \mathcal{A}$ are linear, then there exists a unique linear map $\varphi \times \psi : \mathcal{X} \otimes \mathcal{Y} \to \mathcal{A}$ such that $(\varphi \times \psi)(x \otimes y) = \varphi(x)\psi(y)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

*When $\mathcal{X}, \mathcal{Y}$ are algebras and $\varphi, \psi$ are homomorphisms with commuting ranges, $\varphi \times \psi$ is a homomorphism.*

*Proof.* Exercise 13.1.6. □

When $\mathcal{A}, \mathcal{B}$ are C*-algebras and $\varphi \in \mathcal{A}^*, \psi \in \mathcal{B}^*$, the maps $\varphi \otimes \psi$ and $\varphi \times \psi$ agree via the natural identification $\mathbb{C} \otimes \mathbb{C} = \mathbb{C}$.

It is important to remark that in the situation of Corollary 13.1.9, particularly when $\mathcal{X}, \mathcal{Y}$ are algebras and $\varphi, \psi$ homomorphisms, it is possible to have both of these injective and still $\varphi \times \psi$ not injective; this problem already shows up in dimension 2 (Exercise 13.1.7).

*13.1.10. Remark.* While we have "only" constructed the tensor product of vector spaces, the reader should keep in mind that many objects are vector spaces: Hilbert spaces, Banach spaces, and C*-algebras. So our definition includes algebraic tensor products of C*-algebras, or even of a C*-algebra and a Hilbert space, or of a Banach space and a C*-algebra, etc. There is a caveat, though, which is that one might need to define additional structure, like the inner product for tensor products of Hilbert spaces, or the multiplicative structure and the norm for tensor products of C*-algebras; and, in these cases, we also want to take the completion. Because of that we will only consider as new categories those two kinds of tensor products, and later von Neumann algebras. Other tensor products are very interesting, like those of Banach spaces, but we will not consider them here. Famously, Grothendieck's thesis [Gro55] is a masterpiece in the subject.

We collect here a result that will be needed later.

***13.1.11. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a factor. Then the map $\rho : T \otimes S \longmapsto TS$ induces a $*$-isomorphism from $\mathcal{M} \otimes \mathcal{M}'$ onto* $\operatorname{alg}\{\mathcal{M} \cup \mathcal{M}'\}$.

*Proof.* It is entirely straightforward to check that $\rho$ is a $*$-homomorphism. What is not clear (and not true without the hypotheses) is that $\rho$ is well-defined and injective (see Exercise 13.1.7 for an example of how $\rho$ can fail to be injective). The fact that $\rho$ is well-defined follows from Proposition 13.1.4; indeed, if $\sum_j T_j \otimes S_j = 0$, there exist scalars such that $\sum_k c_{kj} T_k = 0$ and $\sum_j c_{kj} S_j = S_k$. Then

$$\rho\Big(\sum_k T_k \otimes S_k\Big) = \sum_k T_k S_k = \sum_k \sum_j c_{kj} T_k S_j = \sum_j \Big(\sum_k c_{kj} T_k\Big) S_j = 0.$$

So we need to check that $\rho$ is injective. Suppose that $\displaystyle\sum_{j=1}^n T_j S_j = 0$. Let $\mathcal{X} \subset \mathcal{H} \otimes \mathbb{C}^n$ be

$$\mathcal{X} = \overline{\operatorname{span}}\Big\{\sum_{j=1}^n R S_j \xi \otimes e_j : \ R \in \mathcal{M}', \ \xi \in \mathcal{H}\Big\},$$

and let $Q \in \mathcal{B}(\mathcal{H} \otimes \mathbb{C}^n)$ be the orthogonal projection onto $\mathcal{X}$. We may think $Q \in M_n(\mathcal{B}(\mathcal{H}))$, that is

$$Q = \sum_{k,j=1}^{n} Q_{kj} \otimes E_{kj}.$$

For any $R \in \mathcal{M}'$ and $\xi \in \mathcal{H}$,

$$\sum_{j=1}^{n} RS_j\xi \otimes e_j = Q \sum_{j=1}^{n} RS_j\xi \otimes e_j = \sum_{k,j=1}^{n} Q_{kj}RS_j\xi \otimes e_k.$$

As $\{e_1, \ldots, e_n\}$ are linearly independent and $\xi$ was arbitrary , we get that

$$RS_k = \sum_{j=1}^{n} Q_{kj}RS_j, \qquad\qquad k = 1, \ldots, n,\ R \in \mathcal{M}'. \tag{13.1}$$

The hypothesis $\sum_j T_jS_j = 0$ implies that $\sum_{j=1}^n T_j^*\eta \otimes e_j \in \mathcal{X}^\perp$ for all $\eta \in \mathcal{H}$; indeed, since $R \in \mathcal{M}'$

$$\begin{aligned}\Big\langle \sum_{j=1}^{n} RS_j\xi \otimes e_j, \sum_{j=1}^{n} T_j^*\eta \otimes e_j \Big\rangle &= \sum_{k,j=1}^{n} \langle RS_j\xi, T_k^*\eta\rangle \, \langle e_j, e_k\rangle = \sum_{j=1}^{n} \langle T_jS_j\xi, R^*\eta\rangle \\ &= \Big\langle \sum_{j=1}^{n} T_jS_j\xi, R^*\eta \Big\rangle = 0.\end{aligned}$$

This means that

$$0 = Q\sum_{j=1}^{n} T_j^*\eta \otimes e_j = \sum_{k,j=1}^{n} Q_{kj}T_j^*\eta \otimes e_k$$

for all $\eta \in \mathcal{H}$, and again by the linear independence $\displaystyle\sum_{j=1}^{n} Q_{kj}T_j^* = 0$. Taking adjoints (and noting that $Q = Q^*$, so $Q_{kj}^* = Q_{jk}$),

$$\sum_{j=1}^{n} T_jQ_{jk} = 0. \tag{13.2}$$

If we show that $Q_{kj} \in \mathbb{C}$ for all $k, j$, then (13.1) and (13.2) give us $\sum_j T_j \otimes S_j = 0$ by Proposition 13.1.4. For any $S \in \mathcal{M}'$ we have $(S \otimes I_n)\mathcal{X} \subset \mathcal{X}$. And for $T \in \mathcal{M}$

$$(T \otimes I_n)\sum_{j=1}^{n} RS_j\xi \otimes e_j = \sum_{j=1}^{n} TRS_j\xi \otimes e_j = \sum_{j=1}^{n} RS_jT\xi \otimes e_j \in \mathcal{X},$$

so $(T \otimes I_n)\mathcal{X} \subset \mathcal{X}$. It follows that $\mathcal{X}$ is invariant under $\operatorname{alg}\{\mathcal{M} \cup \mathcal{M}') \otimes I_n$, and being closed it is also invariant under $\overline{\operatorname{alg}\{\mathcal{M} \cup \mathcal{M}')}^{\text{sot}} \otimes I_n = (\mathcal{M} \cup \mathcal{M}')'' \otimes I_n = \mathcal{B}(\mathcal{H}) \otimes I_n$, the last equality justified by

$$(\mathcal{M} \cup \mathcal{M}')'' = \big((\mathcal{M} \cup \mathcal{M}')'\big)' = (\mathcal{M}' \cap \mathcal{M})' = (\mathbb{C}\, I_h)' = \mathcal{B}(\mathcal{H}).$$

This means that $Q$ commutes with $T \otimes I_n$ for all $T \in \mathcal{B}(\mathcal{H})$. Thus

$$\begin{aligned} Q_{kj}\xi \otimes e_\ell &= (I_{\mathcal{H}} \otimes E_{\ell k})Q(I_{\mathcal{H}} \otimes E_{j\ell})(\xi \otimes e_\ell) = (I_{\mathcal{H}} \otimes E_{\ell k})Q(\xi \otimes e_j) \\ &= (I_{\mathcal{H}} \otimes E_{\ell k})Q(P_\xi \otimes I_n)(\xi \otimes e_j) = (I_{\mathcal{H}} \otimes E_{\ell k})(P_\xi \otimes I_n)Q(\xi \otimes e_j) \\ &= (P_\xi \otimes E_{\ell k})Q(\xi \otimes e_j) = \lambda_{kj}\xi \otimes e_\ell \end{aligned}$$

for $\lambda_{kj} \in \mathbb{C}$ since $P_\xi \otimes E_{\ell k}$ is rank-one. This occurs for all $\ell$, so $Q_{kj}\xi = \lambda_{kj}\xi$, and $Q_{kj}$ is a scalar. $\square$

**13.1.1. Exercises.**

(13.1.1) Prove Proposition 13.1.2.

(13.1.2) Let $\{e_k\}$ be a basis for $\mathcal{X}$ and $\{f_j\}$ be a basis for $\mathcal{Y}$. Show that $B = \{e_k \otimes f_j\}_{k,j}$ is a basis for $\mathcal{X} \otimes \mathcal{Y}$.

(13.1.3) Let $X, Y$ be complex vector spaces and $X \otimes' Y$ a tensor product defined in some way other than via our bilinear maps. This means that $X \otimes' Y$ is a vector space, spanned by elements of the form $x \otimes' y$ which satisfy the properties in Propositions 13.1.2 and 13.1.3. Show that $X \otimes' Y \simeq X \otimes Y$ canonically.

(13.1.4) Use Theorem 13.1.6 to show that the map $\mathcal{X} \otimes \mathcal{Y} \to \mathcal{Y} \otimes \mathcal{X}$ induced by $x \otimes y \longmapsto y \otimes x$ is an isomorphism.

(13.1.5) Using Theorem 13.1.6 as in Exercise 13.1.4, show that there are canonical isomorphisms (in that they do the obvious thing to the elementary tensors) as follows:

(i) $\mathbb{C} \otimes \mathcal{X} \simeq \mathcal{X}$;

(ii) $\mathbb{C}^n \otimes \mathcal{X} \simeq \mathcal{X}^n$;

(iii) $\mathcal{X} \otimes \mathcal{Y}^* \simeq B(\mathcal{Y}, \mathcal{X})$, if $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional;

(iv) $M_n(\mathbb{C}) \otimes \mathcal{X} \simeq M_n(\mathcal{X})$.

*Note that when $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional one could establish the existence of isomorphisms as above by dimension considerations. But we do not always require finite-dimension, and we want our isomorphisms to be* canonical*.*

(13.1.6) Prove Corollary 13.1.9.

(13.1.7) In the situation of Corollary 13.1.9 where $\mathcal{X}, \mathcal{Y}$ are algebras and $\varphi, \psi$ homomorphisms, show an example where both $\varphi$ and $\psi$ are injective but $\varphi \times \psi$ is not (*Hint: abelian algebras and finite-dimension are enough*).

## 13.2. Completely Positive Maps

We saw in Section 11.5 that we have a natural notion of positivity for linear functionals on a C*-algebra. Complete positivity is the right generalization for linear maps on a C*-algebra where now the codomain is allowed to be a C*-algebra. There is a nice interplay between positivity and ultiplication in the C*-algebra. This will show up in the fact that though completely positive maps are "just" linear, their positivity somehow remembers the algebra multiplication. Completely positive maps are crucial in the treatment of tensor products, and in other parts of the theory of C*-algebras and von Neumann algebras. This will be clearer as we go along.

Since we will cover several different topics related to completely positive maps, and this is a long section, we index the subsections here for easy access:



### 13.2.1. Operator Systems.

**13.2.1. Definition.**

*An* ***operator system*** $\mathcal{S}$ *is a subspace* $\mathcal{S} \subset \mathcal{A}$ *of a unital C*-algebra* $\mathcal{A}$*, that is* ***selfadjoint*** *(that is, it contains the adjoints of all its elements) and* ***unital*** *(that is, it contains the unit of* $\mathcal{A}$*).*

*13.2.2. Remark.* If $\mathcal{S} \subset \mathcal{A}$ is an operator system and $b \in \mathcal{S}$ is positive, then $b = a^*a$ with $a \in \mathcal{A}$, but we usually don't expect $a \in \mathcal{S}$. Still, every operator system has many positive elements: given $a \in \mathcal{S}$, we always have $\operatorname{Re} a, \operatorname{Im} a \in \mathcal{S}$, and hence the positive element $\operatorname{Re} a + \|a\| \, I_{\mathcal{A}}$ is in $\mathcal{S}$. It follows that the positive cone $\mathcal{S}^+$ spans all of $\mathcal{S}$. Explicitly, if $a \in \mathcal{S}$ then

$$a = \big(\operatorname{Re} a + \|a\| \, I_{\mathcal{A}}\big) - \|a\| \, I_{\mathcal{A}} + i\big(\operatorname{Im} a + \|a\| \, I_{\mathcal{A}}\big) - i\|a\| \, I_{\mathcal{A}}.$$

*13.2.3. Examples.* Here are a few examples of operator systems. As we defined operator systems as certain subspaces of a unital C*-algebra, we need to mention C*-algebras too.

(1) Every unital C*-algebra is an operator system. So, for instance, $\mathcal{B}(\mathcal{H})$, $M_n(\mathbb{C})$ and $C(X)$ for any compact Hausdorff space are operator systems.
(2) $\mathbb{C}[x] \subset C[0,1]$ is an operator system that is not a C*-algebra. It is a $*$-algebra, though.
(3) $\mathcal{S} = \{p \in \mathbb{C}[x] : \ \deg p \leq 5\} \subset C[0,1]$ is an operator system that is not an algebra. Of course, 5 can be replaced by any other natural number.
(4)
$$\mathcal{S} = \left\{ \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix} : \ \alpha, \beta, \gamma \in \mathbb{C} \right\} \subset M_2(\mathbb{C})$$
is not an operator system, because $E_{12} \in \mathcal{S}$ but $E_{12}^* = E_{21} \notin \mathcal{S}$.
(5)
$$\mathcal{S} = \left\{ \begin{bmatrix} \alpha & \beta \\ \gamma & \alpha \end{bmatrix} : \ \alpha, \beta, \gamma \in \mathbb{C} \right\} \subset M_2(\mathbb{C})$$
is an operator system that is not an algebra.
(6) $\mathcal{S} = \overline{\operatorname{span}}^{\|\cdot\|} \{1, z, \bar{z}\} \subset C(\mathbb{T})$ is an operator system.

**13.2.2. Completely Positive Maps.** The morphisms for operator systems are the unital *completely positive* maps. Complete positivity appears naturally in many settings, most notably quantum physics and operator algebras.

It is easy to show that $M_n(\mathcal{S}) \subset M_n(\mathcal{A})$ is an operator system (Exercise 13.2.1), since $M_n(\mathcal{A})$ is a C*-algebra. To work on block matrices we will use the notation $a \otimes E_{kj}$ to indicate the matrix with $a$ in the $k, j$ entry and zeros elsewhere; we can see this via Exercise 13.1.5, since $M_n(\mathcal{S}) \simeq \mathcal{S} \otimes M_n(\mathbb{C})$ canonically.

**13.2.4. Definition.**

*Given operator systems $\mathcal{S}, \mathcal{T}$ and $\phi : \mathcal{S} \to \mathcal{T}$, linear, we say that $\phi$ is* ***positive*** *if $\phi(\mathcal{S}^+) \subset \mathcal{T}^+$. We denote by $\phi^{(n)}$ the $n^{\text{th}}$* ***amplification*** *of*

*$\phi$; this is the linear map $\phi^{(n)} : M_n(\mathcal{S}) \to M_n(\mathcal{T})$ induced by*

$$\phi^{(n)}(a \otimes E_{kj}) = \phi(a) \otimes E_{kj}. \tag{13.3}$$

*We say that $\phi$ is $n$-**positive** if $\phi^{(n)}$ is positive, and **completely positive** if it is $n$-positive for all $n$. We sometimes write **cp** in lieu of completely positive, and **ucp** in lieu of unital completely positive.*

In other words,

$$\phi^{(n)}\left(\begin{bmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & & \ddots \end{bmatrix}\right) = \begin{bmatrix} \phi(x_{11}) & \phi(x_{12}) & \cdots \\ \phi(x_{21}) & \phi(x_{22}) & \cdots \\ \vdots & & \ddots \end{bmatrix}.$$

In the context of Corollary 13.1.8 we can write $\phi^{(n)} = \phi \otimes I_n$. It is not obvious at all why this definition of amplification leads to anything useful and/or nice; but we will find out in due course.

A natural notion that comes with amplifications is that of **complete boundedness**.

**13.2.5. Definition.**

*Given $\phi : \mathcal{S} \to \mathcal{T}$ linear, its **completely bounded norm** is given by*

$$\|\phi\|_{\mathrm{cb}} = \sup\{\|\phi^{(n)}\| : \ n \in \mathbb{N}\}.$$

*When $\|\phi\|_{\mathrm{cb}} < \infty$, we say that $\phi$ is **completely bounded**. If $\|\phi\|_{\mathrm{cb}} \leq 1$, we say that $\phi$ is **completely contractive**, or a **complete contraction**.*

Among some very basic facts about amplifications, we have that if $n, m \in \mathbb{N}$ with $n < m$, and $\phi : \mathcal{S} \to \mathcal{T}$, then $\|\phi^{(m)}\| \geq \|\phi^{(n)}\|$, and $\phi^{(m)} \geq 0$ implies $\phi^{(n)} \geq 0$ (Exercise 13.2.2).

*13.2.6. Examples.* Below are some easy examples of completely positive maps.

(1) Let $\mathcal{S} = \mathcal{A}$, $\mathcal{T} = \mathcal{B}$ be C*-algebras, and $\phi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Then $\phi$ is completely positive (Exercise 13.2.3).

(2) Let $\mathcal{S} = \mathcal{A}$ be a C*-algebra, and $b \in \mathcal{A}$. Then $\phi : \mathcal{A} \to \mathcal{A}$ given by $\phi(a) = bab^*$ is completely positive. Indeed, if $X \in M_n(\mathcal{A})^+$, then

$$\phi^{(n)}(X) = \begin{bmatrix} bx_{11}b^* & bx_{12}b^* & \cdots \\ bx_{21}b^* & bx_{22}b^* & \cdots \\ \vdots & & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} b & 0 & 0 & \cdots \\ 0 & b & 0 & \cdots \\ \vdots & & \ddots & \end{bmatrix} X \begin{bmatrix} b & 0 & 0 & \cdots \\ 0 & b & 0 & \cdots \\ \vdots & & \ddots & \end{bmatrix}^* \geq 0$$

(3) Let $T$ be a compact Hausdorff space and let $\mathcal{S} \subset C(T)$ be some operator system. Here we can take advantage of a useful identification, which is that of $M_n(C(T))$ with $C(T, M_n(\mathbb{C}))$ (Exercise 13.2.5). Fix $t_0 \in T$. Define $\phi : \mathcal{S} \to \mathbb{C}$ by $\phi(f) = f(t_0)$. Then $\phi$ is completely positive. Indeed, if $X \in M_n(\mathcal{S})^+$, then $X = Y^*Y$ for some $Y \in M_n(C(T))$, so $X(t) = Y(t)^*Y(t) \in M_n(\mathbb{C})^+$ for each $t \in T$. As $\phi^{(n)}(X) = X(t_0) \in M_n(\mathbb{C})^+$, $\phi$ is completely positive.

*13.2.7. Remark*. Since any operator system $\mathcal{T}$ can be thought as represented faithfully on some $\mathcal{B}(\mathcal{H})$, any map $\phi : \mathcal{S} \to \mathcal{T}$ can be considered as $\phi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$. So it doesn't really hurt generality to consider $\mathcal{B}(\mathcal{H})$ as the codomain, with the only exception of cases where surjectivity is an issue. Also because of this we will mostly use $I$ to denote the identity as it could be either $I_{\mathcal{A}}$ or $I_{\mathcal{H}}$.

While an operator system is given as a subspace of a C$^*$-algebra, most often, like in the statement of the lemma below, we will just refer to the operator system $\mathcal{S}$ and just imply the C$^*$-algebra $\mathcal{A}$.

Let us first see that positive maps are always bounded and preserve adjoints.

***13.2.8. Lemma.***

*Let $\phi : \mathcal{S} \to \mathcal{T}$ be a linear positive map. Then $\phi$ is bounded, and $\|\phi\| \leq 2\|\phi(I)\|$. For any $a \in \mathcal{S}$, $\phi(a^*) = \phi(a)^*$.*

*Proof.* Let $a \in \mathcal{S}$ be selfadjoint. Since $\phi$ is positive, $\phi(a)$ is selfadjoint (Exercise 13.2.6). We have $-a + \|a\|\, I \geq 0$, and so $0 \leq -\phi(a) + \|a\|\, \phi(I)$. Thus, $\phi(a) \leq \|a\|\, \phi(I) \leq \|a\|\, \|\phi(I)\|\, I$. Doing the same for $-a$, we get that

$$-\|a\|\, \|\phi(I)\|\, I \leq \phi(a) \leq \|a\|\, \|\phi(I)\| I.$$

Then $\sigma(\phi(a)) \subset [-\|a\|\, \|\phi(I)\|, \|a\|\, \|\phi(I)\|]$, so $\|\phi(a)\| \leq \|a\|\, \|\phi(I)\|$.

For arbitrary $a$, we write $a = b + ic$ with $b, c$ selfadjoint. Then

$$\|\phi(a)\| \leq \|\phi(b)\| + \|\phi(c)\| \leq (\|b\| + \|c\|)\, \|\phi(I)\| \leq 2\|a\|\, \|\phi(I)\|.$$

Regarding adjoints, if $a \in \mathcal{S}$ is such that $a = a^*$, then $a = a_1 - a_2$ for $a_1, a_2 \in \mathcal{S}$ positive (we showed this in Remark 13.2.2). Then $\phi(a) = \phi(a_1) - \phi(a_2)$ is selfadjoint, as it is a difference of positives. For arbitrary

$a \in \mathcal{S}$, we have $a = b + ic$ with $b, c \in \mathcal{S}$ selfadjoint. Then

$$\phi(a^*) = \phi(b - ic) = \phi(b) - i\phi(c) = (\phi(b) + i\phi(c))^* = \phi(a)^*,$$

since $\phi(b)$ and $\phi(c)$ are selfadjoint. □

*13.2.9. Example.* The estimate from Lemma 13.2.8 looks crude, but it isn't. The following example, due to Arveson, is taken from [Pau02]. Let $\mathcal{S} = \overline{\operatorname{span}}^{\|\cdot\|}\{1, z, \bar{z}\} \subset C(\overline{\mathbb{D}})$, $\mathcal{B} = M_2(\mathbb{C})$, and $\phi : \mathcal{S} \to \mathcal{B}$ be given by

$$\phi(\alpha + \beta z + \gamma \bar{z}) = \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix}$$

One can check (Exercise 13.2.7) that

$$\alpha + \beta z + \gamma \bar{z} \geq 0,\ z \in \overline{\mathbb{D}} \iff \alpha \geq 0,\ \gamma = \bar{\beta}, 2|\beta| \leq \alpha \iff \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix} \geq 0.$$

Then the map $\phi(\alpha + \beta z + \gamma \bar{z}) = \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix}$ is positive, unital, and

$$\|\phi(z)\| = \left\| \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \right\| = 2 = 2\|\phi(1)\|,$$

so $\|\phi\| = 2\|\phi(1)\|$.

Refer to Example 13.2.20 for a different example. The map $\phi$ cannot be completely positive in light of Proposition 13.2.24.

The following is a very useful criterion to check complete positivity.

***13.2.10. Lemma.***

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a linear map. Then the following statements are equivalent:*

*(1) $\phi$ is $n$-positive;*

*(2) for every $a_1, \ldots, a_n \in \mathcal{A}$ and $\xi_1, \ldots, \xi_n \in \mathcal{H}$,*

$$\sum_{k,j=1}^{n} \langle \phi(a_j^* a_k)\xi_k, \xi_j \rangle \geq 0.$$

*Proof.* The required observation is that

$$\sum_{k,j=1}^{n} \langle \phi(a_j^* a_k)\xi_k, \xi_j \rangle = \left\langle \phi^{(n)}(A^*A)\xi, \xi \right\rangle, \tag{13.4}$$

where $\xi = (\xi_1, \ldots, \xi_n)^\top \in \mathcal{H}^n$ and $A \in M_n(\mathcal{A})$ is the matrix with some row $a_1, \ldots, a_n$ and zeroes elsewhere (Exercise 13.2.8). It follows from (13.4) that if $\phi$ is $n$-positive, then the left-hand-side in (13.4) is positive.

Conversely, if the left-hand-side in (13.4) is positive for all choices of elements $a_1, \ldots, a_n$ and vectors $\xi_1, \ldots, \xi_n$, we want to deduce that $\phi$ is $n$-positive. Let $X \in M_n(\mathcal{A})$. We can write $X = X_1 + \cdots + X_n$, where $X_k$ is the matrix such that its $k^{\text{th}}$ row is that of $X$, and the rest of the rows are zero. It is then straightforward to check that $X^*X = \sum_k X_k^* X_k$ (Exercise 13.2.9), and we have

$$\phi^{(n)}(X^*X) = \sum_k \phi^{(n)}(X_k^* X_k),$$

which is a sum of positives—thus positive—by (13.4). $\square$

For an operator system $\mathcal{S}$ we can consider its state space $S(\mathcal{S})$. And since each operator system is embedded in a C*-algebra, it follows that elements in an operator system are also separated by its states, and that the norm and positivity of an element can be characterized by states, as in Proposition 11.5.4 and Corollary 11.5.8 (we are in the unital case here, so the condition (11.12) becomes $\|\varphi\| = \varphi(I_{\mathcal{A}})$).

The statements of the theorems and propositions below will be sometimes stated for operator systems, and sometimes for C*-algebras. This is not arbitrary, but it is due to the fact that we need to tread carefully to prove stuff about operator systems. In many cases we will initially prove a result for C*-algebras and later we will show that the same result holds for operator systems. Krein's Theorem 13.2.11 allows us to extend positive functionals from an operator system to a C*-algebra; for completely positive maps we will have to wait until Arveson's Extension Theorem (13.2.59).

We recall from Proposition 11.5.4 that for $f : \mathcal{A} \to \mathbb{C}$ linear, $f \geq 0$ if and only if $\|f\| = f(I_{\mathcal{A}})$. In Theorem 11.5.7 we used this to extend states from a C*-subalgebra. A quick analysis of the proof shows that replacing $\mathcal{B}$ with an operator system $\mathcal{S}$ changes nothing, since $I_{\mathcal{A}} \in \mathcal{S}$. We state this as

***13.2.11. Theorem. (Krein)***

*Let $\mathcal{S} \subset \mathcal{A}$ be an operator system and $f : \mathcal{S} \to \mathbb{C}$ a positive linear functional. Then there exists $\tilde{f} : \mathcal{A} \to \mathbb{C}$, linear and positive, extending $f$.*

We follow with some useful consequences.

**13.2.12. Proposition.**

*Let $f : \mathcal{S} \to \mathbb{C}$ be a positive linear functional. Then $f$ is completely positive.*

*Proof.* By Theorem 13.2.11 we may assume that $f \in S(\mathcal{A})$. Consider the amplification $f^{(n)} : M_n(\mathcal{A}) \to M_n(\mathbb{C})$. Fix $A \in M_n(\mathcal{A})^+$ and $\xi = (c_1, \ldots, c_n)^\top \in \mathbb{C}^n$—this is the key fact here, that $M_n(\mathbb{C})$ acts on $\mathbb{C}^n$. Let

$$X = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ c_2 & 0 & \cdots & 0 \\ & & \cdots & \\ c_n & 0 & \cdots & 0 \end{bmatrix} \in M_n(\mathbb{C}).$$

Then $\|X\| = \|\xi\|$ (Exercise 13.2.10). As $X^*AX \in M_n(\mathcal{A})^+$,

$$0 \le (X^*AX)_{11} = \sum_{k=1}^n \sum_{j=1}^n a_{kj} c_j \bar{c}_k \in \mathcal{A},$$

and then

$$\begin{aligned} \langle f^{(n)}(A)\xi, \xi\rangle &= \sum_{k=1}^n \sum_{j=1}^n f(a_{kj}) c_j \bar{c}_k = f\left(\sum_{k=1}^n \sum_{j=1}^n a_{kj} c_j \bar{c}_k\right) \\ &= f\left((X^*AX)_{11}\right) \ge 0. \end{aligned}$$

As $\xi \in \mathbb{C}^n$ is arbitrary, we conclude that $f^{(n)}(A) \ge 0$, and so $f$ is completely positive. □

Proposition 13.2.12 is a particular case of Proposition 13.2.22 (but it is used in its proof).

*13.2.13. Remark.* For an operator system $\mathcal{S}$, $M_n(\mathcal{S})$ is also an operator system. For $A, B \in M_n(\mathcal{S})$, the product $AB$ makes in general no sense in $M_n(\mathcal{S})$. But in the case where one of them is a scalar matrix, say $A \in M_n(\mathbb{C})$, then $AB \in M_n(\mathcal{S})$ (Exercise 13.2.11).

**13.2.14. Proposition.**

*Let $\mathcal{S}$ be an operator system, and fix $a \in \mathcal{S}$. Then*

$$\|a\| = \sup\{f(a^*a)^{1/2} : \ f \in S(\mathcal{S})\}. \tag{13.5}$$

*Moreover,*

$$\mathcal{S}^+ = \{a \in \mathcal{S} : \ f(a) \geq 0 \ \forall f \in S(\mathcal{S})\}, \tag{13.6}$$

*and if* $a \in \mathcal{S}^+$ *then* $\|a\| = \max\{f(a) : \ f \in S(\mathcal{S})\}$.

*Proof.* Since $\mathcal{S} \subset \mathcal{A}$ for some C*-algebra $\mathcal{A}$, $\mathcal{S}^+ = \mathcal{A}^+ \cap \mathcal{S}$, and every state on $\mathcal{S}$ is the restriction of a state on $\mathcal{A}$ by Theorem 13.2.11, we will show that the equalities above hold for all elements in $\mathcal{A}$.

For all $a \in \mathcal{A}$, $f \in S(\mathcal{A})$, we have $f(a^*a) \leq \|a^*a\| = \|a\|^2$. Conversely, we can always consider $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$. Fix $\varepsilon > 0$ and choose $\xi \in \mathcal{H}$ a unit vector with $\|a\xi\| > \|a\| - \varepsilon$. The linear map $f(b) = \langle b\xi, \xi\rangle$ is a state, and $f(a^*a) = \|a\xi\|^2 > (\|a\| - \varepsilon)^2$. As $\xi$ and $\varepsilon$ were arbitrary, this establishes (13.5).

If $f(a) \geq 0$ for all states, we have in particular that $\langle a\xi, \xi\rangle \geq 0$ for all $\xi \in H$, and so $a \in \mathcal{A}^+$ and this gives (13.6). Finally, for the last equality, if $a \in \mathcal{A}^+$ then Corollary 11.5.8 provides us with $f \in S(\mathcal{S})$ and $\|a\| = f(a)$. ☐

**13.2.3. Stinespring's Dilation.** If one looks carefully at the examples of completely positive maps, they are all made of combinations of $*$-homomorphisms and conjugation by an element and its adjoint. This is not due to lack of creativity, but is actually due to the fact that every completely positive map arises that way.

***13.2.15. Theorem. (Stinespring's Dilation)***

*Let* $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ *be a completely positive map. Then there exists a Hilbert space* $\mathcal{K}$*, a bounded linear map* $V : \mathcal{H} \to \mathcal{K}$*, and a* $*$*-representation* $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{K})$ *such that*

$$\phi(a) = V^*\pi(a)V, \quad a \in \mathcal{A}, \qquad \text{and} \qquad \|V\|^2 = \|\phi\|.$$

*With the requirement* $\mathcal{K} = \overline{\pi(\mathcal{A})V\mathcal{H}}$*, the triple* $(\mathcal{K}, \pi, V)$*—said to be* ***minimal****—is unique up to unitary equivalence.*

*The representation* $\pi$ *may be assumed to be faithful, at the cost of losing the minimality.*

*Proof.* Assume first that $\mathcal{A}$ is unital. The construction is a generalization of the GNS construction (Theorem 11.6.2). Let $\mathcal{K}_0 = \mathcal{A} \otimes \mathcal{H}$. Define a sesquilinear form on $\mathcal{K}_0$ by $\langle a \otimes \xi, b \otimes \eta\rangle_{\mathcal{K}} = \langle \phi(b^*a)\xi, \eta\rangle$ and extend by linearity. The complete positivity of $\phi$ entails the positivity of the form: for any $a_1, \dots, a_m \in \mathcal{A}$ and $\xi_1, \dots, \xi_m \in \mathcal{H}$,

$$\left\langle \sum_k a_k \otimes \xi_k, \sum_k a_k \otimes \xi_k \right\rangle_{\mathcal{K}} = \sum_{k,j} \langle \phi(a_j^* a_k)\xi_k, \xi_j \rangle \geq 0 \tag{13.7}$$

by Lemma 13.2.10. We obtain the Hilbert space $\mathcal{K}$ by taking the quotient of $\mathcal{K}_0$ over the subspace $\mathcal{K}_{00} = \{\bar{\xi} \in \mathcal{K}_0 : \langle \bar{\xi}, \bar{\xi} \rangle_{\mathcal{K}} = 0\}$ (see Exercise 13.2.12) and then taking the completion. We define by linearity

$$V\xi = I_{\mathcal{A}} \otimes \xi, \qquad \pi(a)(b \otimes \xi) = ab \otimes \xi.$$

We first show that $\pi(a)$ is well-defined and bounded. We have

$$\begin{aligned}
\left\| \pi(a) \Big( \sum_j x_j \otimes \xi_j \Big) \right\|^2 = \left\| \sum_j a x_j \otimes \xi_j \right\|^2 &= \sum_{k,j} \langle \phi(x_j^* a^* a x_k) \xi_k, \xi_j \rangle \\
&= \left\langle \phi^{(m)}(X^* \, A^* A X)\, \xi, \xi \right\rangle \leq \|A^* A\| \, \langle \phi^{(m)}(X^* X) \xi, \xi \rangle \\
&= \|A\|^2 \sum_{k,j} \langle \phi(x_j^* x_k) \xi_k, \xi_j \rangle \\
&= \|A\|^2 \left\| \sum x_j \otimes \xi_j \right\|^2
\end{aligned}$$

where

$$X = \begin{bmatrix} x_1 & \cdots & x_n \\ & \vdots & \\ 0 & \cdots & 0 \end{bmatrix}, \quad A = \begin{bmatrix} a & & \\ & \ddots & \\ & & a \end{bmatrix}, \quad \xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}.$$

This shows that if $\sum_j x_j \otimes \xi_j = 0$, then $\pi(a) \left( \sum_j x_j \otimes \xi_j \right) = 0$, implying that $\pi(a)$ is well-defined. Also, it shows that $\|\pi(a)\| \leq \|a\|$ on the linear manifold $\mathcal{K}_0 / K_{00}$. As it is linear and bounded, $\pi(a)$ extends to the completion $\mathcal{K}$ (Proposition 6.1.9). So $\pi(a) \in \mathcal{B}(\mathcal{K})$. Now for any $\xi, \eta \in \mathcal{H}$,

$$\begin{aligned}
\langle V^* \pi(a) V \xi, \eta \rangle = \langle \pi(a) V\xi, V\eta \rangle_{\mathcal{K}} &= \langle \pi(a)(I_{\mathcal{A}} \otimes \xi), I_{\mathcal{A}} \otimes \eta \rangle_K \\
&= \langle a \otimes \xi, I_{\mathcal{A}} \otimes \eta \rangle_K = \langle \phi(a)\xi, \eta \rangle,
\end{aligned}$$

implying $V^* \pi(a) V = \phi(a)$.

We leave the uniqueness and the non-unital case as an exercise (Exercise 13.2.13).

As for the faithfulness of the representation, let $\rho : \mathcal{A} \to \mathcal{B}(\mathcal{K}')$ be a faithful representation. Consider the faithful representation $\pi \oplus \rho : \mathcal{A} \to \mathcal{B}(\mathcal{K}) \oplus \mathcal{B}(\mathcal{K}') \subset \mathcal{B}(\mathcal{K} \oplus \mathcal{K}')$, and $V' : \mathcal{H} \to \mathcal{K} \oplus \mathcal{K}'$ given by $V'\xi = V\xi \oplus 0$. Then $\pi \oplus \rho$ is faithful, and we still have $\phi(a) = (V')^* (\pi \oplus \rho)(a) V'$. □

*13.2.16. Remark* . The name Stinespring's "Dilation" is better understood under the following picture. When $\phi$ is ucp, the map $V$ from the Stinespring Dilation is an isometry. Then we can identify $\mathcal{H}$ with the subspace $V\mathcal{H}$ of $\mathcal{K}$

(when we restrict the codomain of $V$ to its image, it becomes a unitary). So we can consider the Stinespring triple $(\mathcal{K}_1, \pi_1, P_{\mathcal{H}})$, where $\mathcal{K}_1 = \mathcal{H} \oplus (\mathcal{K} \ominus V\mathcal{H})$, and so $\phi$ is the "upper left corner" of a representation: $\phi(a) = P_{\mathcal{H}}\, \pi(a)|_{\mathcal{H}}$.

This how the Stinespring Dilation is most often used. It also proves that cp maps are always corners of $*$-homomorphisms.

*13.2.17. Remark*. The proof of Stinespring's Theorem above requires $\mathcal{A}$ to be an algebra, in order to define the inner product. The decomposition can be applied to a completely positive map on an operator system, though, via Arveson's Extension Theorem (13.2.59 below).

*13.2.18. Remark*. The GNS representation (Theorem 11.6.2) is a particular case of Stinespring's Theorem, in the case where $\mathcal{H} = \mathbb{C}$. Recall that a positive linear functional is completely positive by Proposition 13.2.12.

We saw in Example 13.2.9 that unital positive maps need not be contractive. On the other hand, unital completely positive maps seem to be the "right" generalization of state, as follows:

***13.2.19. Theorem.***

*Let $\phi : \mathcal{A} \to \mathcal{B}$ be linear and unital. Then the following statements are equivalent:*

*(1) $\phi$ is cp;*
*(2) $\|\phi\|_{\mathrm{cb}} = 1$.*

*Proof.* If $\phi$ is completely positive, by Stinespring's Theorem $\phi = V^*\pi(\cdot)V$; then $I = \phi(I) = V^*V$, so $VV^* \leq I$, and

$$\begin{aligned} \|\phi(a)\|^2 &= \|V^*\pi(a^*)VV^*\pi(a)V\| \leq \|V^*\pi(a^*a)V\| \\ &\leq \|\pi(a^*a)\| \leq \|a^*a\| = \|a\|^2. \end{aligned}$$

Thus, $\|\phi\| \leq 1$. As we also have $\|\phi\| \geq \|\phi(I)\| = 1$, so $\|\phi\| = 1$. As we can apply the same reasoning to $\phi^{(n)}$, we deduce that $\|\phi^{(n)}\| = 1$ for all $n$; thus $\|\phi\|_{\mathrm{cb}} = 1$.

Conversely, suppose that $\|\phi\|_{\mathrm{cb}} = 1$. Let $A \in M_n(\mathcal{A})^+$; we want to show that $\phi^{(n)}(A) \geq 0$. Fix $f \in S(M_n(\mathcal{B}))$. Then $f \circ \phi^{(n)} : M_n(\mathcal{A}) \to \mathbb{C}$ is unital and $\|f \circ \phi^{(n)}\| \leq \|\phi^{(n)}\| \leq 1$. By Proposition 11.5.4, $f \circ \phi^{(n)}$ is positive. Then

$f(\phi^{(n)}(A)) \geq 0$ for all $f \in S(\mathcal{B})$. By Proposition 13.2.14, $\phi^{(n)}(A) \geq 0$. So $\phi$ is cp. $\square$

*13.2.20. Example.* The hypotheses in Theorem 13.2.19 are more or less tight. A classical example of a positive, linear, contractive, and unital map that is not a complete contraction (and so not completely positive) is given by the transpose map $\phi_T : M_2(\mathbb{C}) \to M_2(\mathbb{C})$. It is easy to check that $\phi_T$ is unital, positive, and $\|\phi_T\| = 1$. But $\phi_T$ is not cp (actually, it is not 2-positive). Indeed, $\phi_T^{(2)} : M_4(\mathbb{C}) \to M_4(\mathbb{C})$ is given by $\phi_T^{(2)}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a^\top & b^\top \\ c^\top & d^\top \end{bmatrix}$. Then we can consider

$$\phi_T^{(2)}\left(\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where the image of a positive matrix is not positive (Exercise 13.2.16). Theorem 13.2.19 then implies that $\phi_T$ is not completely contractive. Actually, the example above, with the role of the two matrices reversed, shows that $\|\phi_T^{(2)}\| \geq 2$, since the norms of the matrices are 2 and 1 respectively. With a bit more work it can be shown that $\|\phi_T\|_{\mathrm{cb}} = 2$.

*13.2.21. Example.* For infinite-dimensional $\mathcal{H}$, the transpose map $\phi_T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ (with respect to a fixed orthonormal basis) is not completely bounded, since $\|\phi_T^{(n)}\| \geq n$. We can basically repeat the reasoning from Example 13.2.20, but because we have infinitely many matrix units available, we can scale the experiment. Indeed, if $\{E_{k,j}\}$ denote matrix units for $\mathcal{B}(\mathcal{H})$ and $\{F_{k,j}\}$ denote matrix units for $M_n(\mathbb{C})$, we can consider $A = \sum_{k,j=1}^n E_{j,k} \otimes F_{k,j} \in M_n(\mathcal{B}(\mathcal{H}))$. Then

$$\begin{aligned} A^*A &= \sum_{k,j=1}^n \sum_{s,t=1}^n E_{t,s}E_{j,k} \otimes F_{s,t}F_{k,j} = \sum_{k,j=1}^n E_{k,k} \otimes F_{j,j} \\ &= \Big(\sum_k E_{k,k}\Big) \otimes \Big(\sum_j F_{j,j}\Big) = I \otimes I, \end{aligned}$$

so $\|A\| = 1$. On the other hand, if

$$B = \phi_T^{(n)}(A) = \sum_{k,j} E_{k,j} \otimes F_{k,j},$$

it is clear that $B = B^*$, and

$$B^2 = \sum_{k,j=1}^{n} \sum_{s,t=1}^{n} E_{s,t}E_{k,j} \otimes F_{s,t}F_{k,j} = \sum_{k=1}^{n} \sum_{s,j} E_{s,j} \otimes F_{s,j} = nB.$$

Then $B = B^2/n$ is positive, and $\|B\|^2 = \|B^2\| = n\|B\|$, so $n = \|B\| = \|\phi_T^{(n)}(A)\|$, implying that $\|\phi_T^{(n)}\| \geq n$.

When a completely positive map is not unital, it is still true that it is completely bounded (Proposition 13.2.24). We could show this by using Stinespring's Dilation. The limitation we have right now that is that we cannot apply Stinespring when the domain is not a C*-algebra. As already mentioned, this will be eventually solved by Arveson's Extension Theorem (13.2.59).

It turns out that the essential feature in Example 13.2.9 is that the codomain is not abelian. Indeed,

***13.2.22. Proposition.***

*Let $\mathcal{A}$ be a unital abelian C*-algebra, and $\psi : \mathcal{S} \to \mathcal{A}$ and $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ positive linear maps. Then $\psi$ and $\phi$ are completely positive.*

*Proof.* We have $\mathcal{A} = C(X)$ for some compact Hausdorff space $X$. Consider first $\psi : \mathcal{S} \to \mathcal{A}$. Note that for each $s \in \mathcal{S}$, $\psi(s)$ is a function on $X$; so we write $\psi(a)(t)$. For each $t \in X$, consider the state $\psi_t(s) = \psi(s)(t)$. To show that $\psi^{(n)}$ is positive, we need to prove that $\psi^{(n)}(A) \geq 0$ for any $A \in M_n(\mathcal{S})^+$. By Exercise 13.2.5 we can naturally identify $M_n(C(X))$ with $C(X, M_n(\mathbb{C}))$. By Proposition 13.2.12, each $\psi_t$ is completely positive, so $\psi_t^{(n)}(A) \geq 0$ for all $t \in X$, which means that $\psi^{(n)}(A)$ is positive.

Now consider $\phi : C(X) \to \mathcal{B}(\mathcal{H})$. Let $A \in M_n(C(X)) = C(X, M_n(\mathbb{C}))$ be positive. Fix $\varepsilon > 0$. By a compactness argument we can get open sets $V_1, \ldots, V_m \subset X$ with $X = \bigcup V_j$ and positive matrices $A_1, \ldots, A_m \in M_n(\mathbb{C})$ such that $\|A(t) - A_j\| < \varepsilon$ for all $t \in V_j$. Let $g_1, \ldots, g_m$ be a partition of unity subordinate to $V_1, \ldots, V_m$. For any $t \in X$,

$$\|A(t) - \sum_j g_j(t)A_j\| = \|\sum_j g_j(t)\,(A(t) - A_j)\| \leq \sum_j g_j(t)\,\varepsilon = \varepsilon.$$

so $\|A - \sum_j g_j A_j\| < \varepsilon$. That is, $A$ is a limit of matrix functions of said form. The fact that $\phi$ is positive implies that it is bounded (Lemma 13.2.8) and so $\phi^{(n)}$ is bounded by the crude estimate $\|\phi^{(n)}(A)\| \leq \sum_{k,j=1}^{n} \|\phi(a_{kj})\| \leq n^2\|\phi\|\,\|A\|$.

So $\phi^{(n)}(A)$ is a limit of function matrices of the form $\phi^{(n)}(\sum_j g_j A_j)$, and thus it will be positive if those matrices are. It is enough to consider $\phi^{(n)}(gA)$, where $g \in C(X)^+$ and $A \in M_n(\mathbb{C})^+$. Then, using that $\phi(g) \geq 0$,

$$\phi^{(n)}(g\,A) = \phi^{(n)}\Big(\sum_{k,j} g\,a_{kj} \otimes E_{kj}\Big) = \sum_{k,j} \phi(g\,a_{kj}) \otimes E_{kj}$$
$$= \sum_{k,j} \phi(g)\,a_{kj} \otimes E_{kj} = \phi(g)\,A \geq 0. \qquad \square$$

Example 13.2.9 shows that the assumption made in Proposition 13.2.22 requiring that the domain of $\phi$ is a C*-algebra—and not just an operator system, even if it lives within an abelian algebra—is essential.

**13.2.4. Block Matrices and Multiplicative Domain.** Let us put results about block matrices to use. Paulsen showed that matricial positivity in an operator system is enough to recover the norm.

***13.2.23. Proposition. (Paulsen)***

*Let $\mathcal{S}$ be an operator system. Then, for any $X \in M_n(\mathcal{S})$,*

$$\|X\| = \inf\left\{ t > 0 : \begin{bmatrix} t\,I & X \\ X^* & t\,I \end{bmatrix} \geq 0 \text{ in } M_{2n}(\mathcal{S}) \right\}.$$

*Proof.* Using Lemma 11.7.2,

$$\begin{bmatrix} t\,I & X \\ X^* & t\,I \end{bmatrix} = t \begin{bmatrix} I & X/t \\ X^*/t & I \end{bmatrix} \geq 0 \iff \|X/t\| \leq 1 \iff \|X\| \leq t. \qquad \square$$

Not surprisingly, results about block matrices are crucial to deal with complete positivity. This first result is key to the proof of Arveson's Extension Theorem 13.2.59. We note that in Theorem 13.2.19 the domain was a C*-algebra, whereas now it is an operator system.

***13.2.24. Proposition.***

*Let $\phi : \mathcal{S} \to \mathcal{B}$ be completely positive. Then $\|\phi\|_{\mathrm{cb}} = \|\phi\| = \|\phi(I)\|$.*

*Proof.* Since $\|\phi(I)\| \leq \|\phi\| \leq \|\phi\|_{\mathrm{cb}}$, only the inequality $\|\phi\|_{\mathrm{cb}} \leq \|\phi(I)\|$ needs proof.

Let $A \in M_n(\mathcal{S})$ with $\|A\| \leq 1$. Then $A^*A \leq I$, and so by Lemma 11.7.2,

$$\begin{bmatrix} I_n & A \\ A^* & I_n \end{bmatrix} \geq 0.$$

As $\phi$ is completely positive,

$$0 \leq \phi^{(2n)}\left(\begin{bmatrix} I_n & A \\ A^* & I_n \end{bmatrix}\right) = \begin{bmatrix} \phi^{(n)}(I_n) & \phi^{(n)}(A) \\ \phi^{(n)}(A)^* & \phi^{(n)}(I_n) \end{bmatrix}.$$

By Lemma 11.7.3, $\|\phi^{(n)}(A)\| \leq \|\phi^{(n)}(I_n)\| = \|\phi(I)\|$. □

***13.2.25. Proposition. (Kadison's Schwarz Inequality)***

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be 2-positive. Then, for any $a \in \mathcal{A}$,*

$$\phi(a)^*\phi(a) \leq \|\phi(I)\|\, \phi(a^*a).$$

*In particular, if $\phi$ is unital,*

$$\phi(a)^*\phi(a) \leq \phi(a^*a). \tag{13.8}$$

*Proof.* Using Lemma 11.7.2 and the trivial inequalities $\phi(I) \leq \|\phi(I)\|\, I$ and $a^*a \geq a^*a$,

$$\begin{bmatrix} \|\phi(I)\|\, I & \phi(a) \\ \phi(a)^* & \phi(a^*a) \end{bmatrix} \geq \begin{bmatrix} \phi(I) & \phi(a) \\ \phi(a)^* & \phi(a^*a) \end{bmatrix} = \phi^{(2)}\left(\begin{bmatrix} I & a \\ a^* & a^*a \end{bmatrix}\right) \geq 0.$$

Then again Lemma 11.7.2 gives us $\frac{\phi(a)^*}{\|\phi(I)\|}\frac{\phi(a)}{\|\phi(I)\|} \leq \frac{\phi(a^*a)}{\|\phi(I)\|}$, so

$$\phi(a)^*\phi(a) \leq \|\phi(I)\|\, \phi(a^*a).$$ □

*13.2.26. Remark* . The inequality (13.8) is most often used for completely positive maps (as one seldom works with 2-positive maps that are not completely positive). In that case one can obtain a rather direct proof from the Stinespring's Dilation. Indeed, if $\phi = V^*\pi V$, then

$$\begin{aligned}\phi(a)^*\phi(a) &= V^*\pi(a^*)VV^*\pi(a)V \leq \|V\|^2\, V^*\pi(a^*)\pi(a)V \\ &= \|V\|^2\, V^*\pi(a^*a)V = \|\phi(I)\|\, \phi(a^*a).\end{aligned}$$

*13.2.27. Remark.* If in Proposition 13.2.25 one restricts the domain to normal elements, then only positivity of $\phi$ is required. Indeed, if $a \in \mathcal{A}$ is normal, $C^*(a)$ is abelian; by Proposition 13.2.22, $\phi$ is cp when restricted to $C^*(a)$.

*13.2.28. Remark.* There are fairly simple examples of positive maps that satisfy Kadison's inequality and are not even 2-positive. For instance let $\phi : M_2(\mathbb{C}) \to$

$M_2(\mathbb{C})$, be given by $\phi(A) = \frac{1}{2}A^\top + \frac{1}{4}\operatorname{Tr}(A)\, I_2$. Then $\phi$ is positive but

$$\phi^{(2)}\left(\begin{bmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{bmatrix}\right) = \begin{bmatrix}3/4&0&0&0\\0&1/4&1/2&0\\0&1/2&1/4&0\\0&0&0&3/4\end{bmatrix}$$

is not positive. Meanwhile, let $\psi(A) = \phi(A^*A) - \phi(A)^*\phi(A)$, and note that $\psi(A+\lambda I) = \psi(A)$ for all $\lambda \in \mathbb{C}$. In particular, taking $B = A - \operatorname{Tr}(A)\, I_2$, we have that $\psi(B) = \psi(A)$ and $\operatorname{Tr}(B) = 0$. Then

$$\begin{aligned}\psi(A) = \psi(B) &= \frac{1}{4}\, B^*B + \frac{1}{2}\, \operatorname{Tr}(B^*B)\, I_2 - \frac{1}{2}\, B^{*\top}\, \frac{1}{2}\, B^\top \\ &= \frac{1}{4}\, B^*B + \frac{1}{2}\, \operatorname{Tr}(B^*B)\, I_2 - \frac{1}{4}\, BB^* \\ &\geq \frac{1}{4} B^*B \geq 0,\end{aligned}$$

since $BB^* \leq \operatorname{Tr}(BB^*)\, I_2 = \operatorname{Tr}(B^*B)\, I_2$. So $\phi$ satisfies Kadison's inequality even though it is not 2-positive.

Using Kadison's inequality we can get a version of Proposition 13.2.24, where we require the domain to be a C$^*$-algebra, but we relax complete positivity to 2-positivity (Exercise 13.2.20).

Kadison's inequality allows for consideration of the **multiplicative domain** of a ucp map $\phi$. Indeed, we have:

***13.2.29. Theorem. (Multiplicative Domain)***

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be 2-positive and unital. Then*

$$\{a \in \mathcal{A} :\ \phi(a^*a) = \phi(a)^*\phi(a)\} = \{a \in \mathcal{A} :\ \phi(ba) = \phi(b)\phi(a)\ \forall b \in \mathcal{A}\}$$

*and*

$$\{a \in \mathcal{A} :\ \phi(aa^*) = \phi(a)\phi(a)^*\} = \{a \in \mathcal{A} :\ \phi(ab) = \phi(a)\phi(b)\ \forall b \in \mathcal{A}\}.$$

*Proof.* The inclusions $\supset$ hold trivially. Now fix $a \in \mathcal{A}$ such that $\phi(a^*a) = \phi(a)^*\phi(a)$, $b \in \mathcal{A}$, and $t \in \mathbb{R}$. Then, using Kadison's Schwarz inequality (13.8),

$$\phi(tb+a)^*\phi(tb+a) \leq \phi((tb+a)^*(tb+a)).$$

After expanding we get

$$t^2\phi(b)^*\phi(b) + 2t\mathrm{Re}\, \phi(b)^*\phi(a) + \phi(a)^*\phi(a) \leq t^2\phi(b^*b) + 2t\mathrm{Re}\, \phi(b^*a) + \phi(a^*a).$$

As $\phi(a)^*\phi(a) = \phi(a^*a)$, we get after regrouping that

$$0 \leq t^2(\phi(b^*b) - \phi(b)^*\phi(b)) + 2t\mathrm{Re}\,(\phi(b^*a) - \phi(b)^*\phi(a)).$$

For this to hold for all $t \in \mathbb{R}$, we need to have $\operatorname{Re}(\phi(b^*a) - \phi(b)^*\phi(a)) = 0$. Doing the same argument with $ib$, we obtain $\operatorname{Im}(\phi(b^*a) - \phi(b)^*\phi(a)) = 0$, so $\phi(b^*a) - \phi(b)^*\phi(a) = 0$. After replacing $b^*$ with $b$, we get $\phi(ba) = \phi(b)\phi(a)$. □

*13.2.30. Remark* . Here is a nice alternative proof of Theorem 13.2.29 that requires 4-positivity (because we need 2-positivity of $\phi^{(2)}$). Let $X = \begin{bmatrix} b^* & a \\ 0 & 0 \end{bmatrix} \in M_2(\mathcal{A})$. By Kadison's Inequality (13.8),

$$\begin{bmatrix} \phi(bb^*) & \phi(ba) \\ \phi(a^*b^*) & \phi(a^*a) \end{bmatrix} = \phi^{(2)}(X^*X) \geq \phi^{(2)}(X)^*\phi^{(2)}(X)$$
$$= \begin{bmatrix} \phi(b)\phi(b)^* & \phi(b)\phi(a) \\ \phi(a)^*\phi(b)^* & \phi(a)^*\phi(a) \end{bmatrix}.$$

We can rewrite this as

$$0 \leq \begin{bmatrix} \phi(bb^*) - \phi(b)\phi(b)^* & \phi(ba) - \phi(b)\phi(a) \\ \phi(a^*b^*) - \phi(a)^*\phi(b)^* & 0 \end{bmatrix}.$$

By Exercise 13.2.18, the $1,2$ entry is zero and so $\phi(ba) = \phi(b)\phi(a)$. The case where $\phi(aa^*) = \phi(a)\phi(a)^*$ is proven analogously.

**13.2.31. Corollary.**

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be 4-positive and unital. Then the set*

$$M_\phi = \{a \in \mathcal{A} : \ \phi(a^*a) = \phi(a)^*\phi(a) \ \text{ and } \phi(aa^*) = \phi(a)\phi(a)^*\}$$

*is a C*-subalgebra of $\mathcal{A}$, called the* ***multiplicative domain*** *of $\phi$.*

*Proof.* By Theorem 13.2.29,

$$M_\phi = \{a \in \mathcal{A} : \ \phi(ab) = \phi(a)\phi(b), \ \phi(ba) = \phi(b)\phi(a) \ \forall b \in \mathcal{A}\}.$$

It is then straightforward to show that $M_\phi$ is a C*-subalgebra of $\mathcal{A}$. □

We show two simple applications of the multiplicative domain. The first one shows that if two C*-algebras are isomorphic as operator systems, then they are isomorphic as C*-algebras.

**13.2.32. Proposition.**

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras, and $\phi : \mathcal{A} \to \mathcal{B}$ a unital 2-positive linear bijection, with $\phi^{-1}$ also 2-positive. Then $\phi$ is a $*$-isomorphism.*

*Proof.* What is not a given is that $\phi$ is multiplicative. Let $a \in \mathcal{A}$, and put $b = \phi(a)$. Then, using Kadison's inequality (13.8) for both $\phi$ and $\phi^{-1}$, and the positivity of $\phi$,

$$\phi(a)^*\phi(a) \le \phi(a^*a) = \phi(\phi^{-1}(b)^*\phi^{-1}(b)) \le \phi(\phi^{-1}(b^*b)) = b^*b = \phi(a)^*\phi(a).$$

Then $\phi(a)^*\phi(a) = \phi(a^*a)$ and so $a \in M_\phi$. Thus $\phi$ is multiplicative, and it is an algebra $*$-isomorphism. □

***13.2.33. Proposition.***

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be ucp. If $\mathcal{B} \subset \mathcal{A}$ is a $C^*$-subalgebra with the same unit and $\phi(b) = b$ for all $b \in \mathcal{B}$, then $\phi(bac) = b\phi(a)c$ for all $a \in \mathcal{A}$, $b, c \in \mathcal{B}$.*

*Proof.* This follows directly from the fact that $\mathcal{B} \subset M_\phi$ and Theorem 13.2.29. □

### 13.2.5. Pure Completely Positive Maps.

**13.2.34. Definition.**

*Let $\phi_1, \phi_2 : \mathcal{S} \to \mathcal{B}(\mathcal{H})$. We write $\phi_1 \le_{\mathrm{CP}} \phi_2$ if $\phi_2 - \phi_1$ is completely positive.*

*A completely positive map $\phi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$ is **pure** if whenever $\psi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$ is completely positive and $\psi \le_{\mathrm{CP}} \phi$, there exists $t \in [0, 1]$ with $\psi = t\,\phi$.*

***13.2.35. Proposition.***

*Let $\phi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$, completely positive. The following statements are equivalent:*

*(i) $\phi$ is pure;*

*(ii) if $\phi = \phi_1 + \phi_2$ with $\phi_1, \phi_2$ completely positive, then $\phi_1 = t_1\,\phi$, $\phi_2 = t_2\,\phi$ for some $t_1, t_2 \in \mathbb{R}$.*

*Proof.* If $\phi$ is pure and $\phi = \phi_1 + \phi_2$, then $\phi - \phi_1 = \phi_2$ is completely positive, so there exists $t_1 \in \mathbb{R}$ with $\phi_1 = t_1\,\phi$. We can reason similarly for $\phi_2$.

Conversely, if $\phi$ satisfies (ii) and $\psi \le_{\mathrm{CP}} \phi$, then $\phi = \psi + (\phi - \psi)$, with both summands completely positive, Then $\psi = t\,\phi$ for some $t \in [0, 1]$. □

### *13.2.36. Proposition.* *(Størmer)*

*Let $\pi : \mathcal{A} \to \mathcal{B}$ be a $*$-homomorphism. Then $\pi$ is extreme in the set of unital positive maps from $\mathcal{A}$ to $\mathcal{B}$.*

*Proof.* Write $\pi = (\phi_1 + \phi_2)/2$, with $\phi_1, \phi_2$ unital and positive. Let $a \in A$ selfadjoint. Then, using Kadison's inequality for unital positive maps (see Remark 13.2.27),

$$\pi(a)^2 = \frac{(\phi_1(a) + \phi_2(a))^2}{4} = \frac{\phi_1(a)^2 + \phi_2(a)^2}{2} - \frac{(\phi_1(a) - \phi_2(a))^2}{4}$$
$$\leq \frac{\phi_1(a)^2 + \phi_2(a)^2}{2} \leq \frac{\phi_1(a^2) + \phi_2(a^2)}{2} = \pi(a^2).$$

Because $\pi$ is multiplicative, $\pi(a)^2 = \pi(a^2)$. So the inequalities above become equalities. In particular, we obtain

$$(\phi_1(a) - \phi_2(a))^2 = 0.$$

Then $\phi_1(a) - \phi_2(a) = 0$, since it is selfadjoint. As this works for any selfadjoint $a \in A$, it follows that $\phi_1 = \phi_2$. □

### *13.2.37. Lemma.*

*Let $\phi_1, \phi_2 : \mathcal{S} \to \mathcal{B}(\mathcal{H})$ be completely positive, with $\phi_1 \leq_{\mathrm{CP}} \phi_2$. Let $(\pi_j, H_j, V_j)$ be minimal Stinespring triples for $\phi_1$ and $\phi_2$. Then there exists a linear contraction $T : \mathcal{H}_2 \to \mathcal{H}_1$ such that $V_1 = TV_2$ and $T\pi_2(a) = \pi_1(a)T$ for all $a \in \mathcal{A}$.*

*Proof.* Using that $\pi_2(\mathcal{A})V_2\mathcal{H}$ is dense in $\mathcal{H}_2$, we define $T\pi_2(a)V_2\xi = \pi_1(a)V_1\xi$. We first need to check that this is a well defined contraction. This follows from the following inequality: given $\xi_1, \ldots, \xi_m \in H$, $a_1, \ldots, a_m \in \mathcal{A}$,

$$\begin{aligned}
\|\sum_j \pi_1(a_j)V_1\xi_j\|^2 &= \sum_{k,j} \langle V_1^*\pi_1(a_k^*a_j)V_1\xi_j, \xi_k\rangle = \sum_{k,j} \langle \phi_1(a_k^*a_j)\xi_j, \xi_k\rangle \\
&= \langle \phi_1^{(m)}(A^*A)\tilde{\xi}, \tilde{\xi}\rangle \leq \langle \phi_2^{(m)}(A^*A)\tilde{\xi}, \tilde{\xi}\rangle \\
&= \sum_{k,j} \langle \phi_2(a_k^*a_j)\xi_j, \xi_k\rangle = \sum_{k,j} \langle V_2^*\pi_2(a_k^*a_j)V_2\xi_j, \xi_k\rangle \\
&= \|\sum_j \pi_2(a_j)V_2\xi_j\|^2,
\end{aligned}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$ and $\tilde{\xi} = \begin{bmatrix} \xi_1 & \cdots & \xi_m \end{bmatrix}^\top$. We have, for any $\xi \in \mathcal{H}$,

$$TV_2\xi = T\pi_2(I_{\mathcal{A}})V_2\xi = \pi_1(I_{\mathcal{A}})V_1\xi = V_1\xi,$$

so $TV_2 = V_1$. Finally,

$$\begin{aligned}\pi_1(a)T\pi_2(b)V_2\xi &= \pi_1(a)\pi_1(b)V_1\xi = \pi_1(ab)V_1\xi\\ &= T\pi_2(ab)V_2\xi = T\pi_2(a)\pi_2(b)V_2\xi\end{aligned}$$

implies that $\pi_1(a)T = T\pi_2(a)$ for all $a \in \mathcal{A}$. □

We discussed commutants at length in Section 12.3.

***13.2.38. Theorem.***

*Let $\psi, \phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be completely positive and let $(\pi, K, V)$ be a minimal Stinespring triple for $\phi$. The following statements are equivalent:*

*(1) $\psi \leq_{\mathrm{CP}} \phi$;*
*(2) there exists a positive contraction $T \in \pi(\mathcal{A})'$ such that $\psi(a) = V^*T\pi(a)V$, $a \in \mathcal{A}$.*

*Proof.* (1) $\implies$ (2) Let $(\pi_0, K_0, V_0)$ be a minimal Stinespring triple for $\psi$. By Lemma 13.2.37 there exists a linear contraction $S : K \to K_0$ with $V_0 = SV$ and $S\pi(a) = \pi_0(a)S$ for all $a \in \mathcal{A}$. Let $T = S^*S$. Then $T$ is a positive contraction, and for any $a \in \mathcal{A}$,

$$\begin{aligned}T\pi(a) = S^*S\pi(a) &= S^*\pi_0(a)S = (\pi_0(a^*)S)^*S\\ &= (S\pi(a))^*S = \pi(a)S^*S = \pi(a)T.\end{aligned}$$

So $T \in \pi(\mathcal{A})'$. Finally, for any $a \in \mathcal{A}$,

$$\psi(a) = V_0^*\pi_0(a)V_0 = V^*S^*\pi_0(a)SV = V^*S^*S\pi(a)V = V^*T\pi(a)V. \qquad \square$$

(2) $\implies$ (1) We have $\psi(a) = V^*T\pi(a)V$ for all $a \in \mathcal{A}$. If $X \in M_n(\mathcal{A})$ then, with

$$\tilde{V} = \begin{bmatrix} V & & \\ & \ddots & \\ & & V \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} T & & \\ & \ddots & \\ & & T \end{bmatrix},$$

we have $\tilde{T} \in M_n(\pi(\mathcal{A}))'$ and

$$\begin{aligned}\psi^{(n)}(X^*X) &= \tilde{V}^*\tilde{T}\,\pi^{(n)}(X^*X)\,\tilde{V} = \tilde{V}^*\,\pi^{(n)}(X)^*\tilde{T}\pi^{(n)}(X)\,\tilde{V}\\ &\leq \tilde{V}^*\,\pi^{(n)}(X)^*\pi^{(n)}(X)\,\tilde{V} = V^*\,\pi^{(n)}(X^*X)\,\tilde{V} = \phi^{(n)}(X^*X).\end{aligned}$$

So $\psi \leq_{\mathrm{CP}} \phi$.

**13.2.39. Definition.**

*A representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ is **irreducible** if $\pi(\mathcal{A})' = \mathbb{C}\, I_{\mathcal{H}}$.*

***13.2.40. Corollary.***

*Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be nonzero and completely positive, and let $(\pi, K, V)$ be a minimal Stinespring dilation for $\phi$. The following statements are equivalent:*

*(1) $\phi$ is pure;*
*(2) $\pi$ is irreducible.*

*Proof.* Assume first that $\phi$ is pure. Let $T \in \pi(\mathcal{A})'$ with $0 \le T \le I$. Put $\psi(a) = V^*T\pi(a)V$, $a \in \mathcal{A}$. Then $\psi \le_{\mathrm{CP}} \phi$, and so there exists $t \in \mathbb{R}$ with $\psi = t\,\phi$. Thus $V^*T\pi(a)V = t\,V^*\pi(a)V$ for all $a \in \mathcal{A}$. Given $\xi, \eta \in \mathcal{H}$ and $b, c \in \mathcal{A}$,

$$\langle T\,\pi(b)V\xi, \pi(c)V\eta\rangle = \langle V^*T\pi(c^*b)V\xi, \eta\rangle = t\,\langle \pi(b)V\xi, \pi(c)V\eta\rangle,$$

implying that $T\pi(b)V\xi = t\pi(b)V\xi$, and so by density $T = t\,I$. As positive elements span all of $\pi(\mathcal{A})'$ we get that $\pi(\mathcal{A})' = \mathbb{C}\,I$, so $\pi$ is irreducible.

Conversely, if $\pi$ is irreducible and $\psi \le_{\mathrm{CP}} \phi$, by Theorem 13.2.38 there exists a positive contraction $T \in \pi(\mathcal{A})'$ such that $\psi(a) = V^*T\pi(a)V$ for all $a \in \mathcal{A}$. As $\pi$ is irreducible, $\pi(\mathcal{A})' = \mathbb{C}\,I$, so $T = t\,I$ for some $t \in \mathbb{R}$. So $\psi = t\,\phi$, and thus $\phi$ is pure. $\square$

In the case where $\dim \mathcal{H} = 1$ (so $\mathcal{H} = \mathbb{C}$), we get a notion of **pure state**. It turns out that the pure states are precisely the extreme points of the state space $S(\mathcal{A})$:

***13.2.41. Proposition.***

*Let $\mathcal{A}$ be a $C^*$-algebra, $f \in S(\mathcal{A})$. The following statements are equivalent:*

*(1) $f$ is pure;*
*(2) $f$ is extreme on $S(\mathcal{A})$;*
*(3) the GNS representation of $f$ is irreducible.*

*Proof.* Assume first that $f$ is pure. If $f = tf_1 + (1-t)f_2$ for states $f_1, f_2 \in S(\mathcal{A})$ and $t \in [0,1]$, then $tf_1 \le f$, and so $tf_1 = rf$ for some $r \in [0,1]$; evaluating at $I$, we get $r = t$, and so $f_1 = f$. It follows immediately that $f_2 = f$, and so $f$ is extreme.

Conversely, suppose that $f$ is extreme and $g \le f$ for some positive functional $g$; if $g$ is a state, then $f - g$ is a positive functional with $(f-g)(I) = 0$, and so $f = g$ by Proposition 11.5.4. If $g(I) = 0$ then $g = 0$. So we may assume that $0 < g(I) < 1$; in that case,

$$f(a) = g(I)\,\frac{g(a)}{g(I)} + (1-g(I))\,\frac{f(a)-g(a)}{1-g(I)}, \qquad a \in \mathcal{A},$$

a convex combination of states. As $f$ is extreme, $\frac{g}{g(I)} = f$, and we conclude that $f$ is pure.

The equivalence with the irreducibility of the GNS representation requires noticing that, for a state, the GNS representation and the Stinespring dilation yield the same construction. After that, the equivalence was already proven in Corollary 13.2.40. □

***13.2.42. Corollary.***

*A C*-algebra has lots of pure states. Concretely, the pure states of a C*-algebra separate points.*

*Proof.* Let $\mathcal{A}$ be a C*-algebra. Since we want to talk about pure states, by Proposition 11.5.6 and Exercise 13.2.21 we may assume without loss of generality that $\mathcal{A}$ is unital. By Proposition 11.5.15, $S(\mathcal{A})$ is weak*-compact. By Krein–Milman (Theorem 7.5.11) and Proposition 13.2.41,

$$S(\mathcal{A}) = \overline{\operatorname{conv}}^{w^*} S_P(\mathcal{A}), \tag{13.9}$$

where $S_P(\mathcal{A})$ denotes the set of pure states. It follows that if $\varphi(a) = 0$ for all pure states $\varphi$, then $\varphi(a) = 0$ for all $\varphi \in S(\mathcal{A})$; by Corollary 11.5.8, $a = 0$. □

***13.2.43. Corollary.***

*Let $\mathcal{A}$ be a C*-algebra such that all its irreducible representations are faithful. Then $\mathcal{A}$ is simple.*

*Proof.* Let $\mathcal{J} \subset \mathcal{A}$ be a proper ideal. By Corollary 13.2.42 there exists a pure state $\varphi$ on $\mathcal{A}/\mathcal{J}$. Then Proposition 13.2.41 implies that the GNS representation $\pi_\varphi$ is irreducible. If $q : \mathcal{A} \to \mathcal{A}/\mathcal{J}$ denotes the quotient map, it follows that $\pi_\varphi \circ q : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\varphi)$ is an irreducible representation, which is faithful by hypothesis. If $a \in J$ we have $q(a) = 0$, so $(\pi_\varphi \circ q)(a) = 0$; by the faithfulness, $a = 0$, and hence $\mathcal{J} = \{0\}$. So $\mathcal{A}$ is simple. □

For a state, being pure is kind of the opposite of being faithful. Concretely,

***13.2.44. Lemma.***

*Let $\mathcal{A}$ be a C*-algebra and $\varphi \in S(\mathcal{A})$ pure and faithful. Then $\mathcal{A} = \mathbb{C}$.*

*Proof.* Assume first that $\mathcal{A}$ is abelian. Then $\mathcal{A} \simeq C_0(X)$ for some locally compact Hausdorff space $X$. We know from Example (7.5.10) and Proposition 13.2.41 that the pure states are precisely the point evaluations. So there exists $x_0 \in X$ such that $\varphi(g) = g(x_0)$ for all $g \in C_0(X)$. For any $g \in C_0(X)$ we have $\varphi(g - g(x_0)) = 0$; as $\varphi$ is faithful, $g = g(x_0)$ is constant. If every continuous function is constant, then $X = \{x_0\}$ (Exercise 2.6.8). So $\mathcal{A} \simeq C_0(\{x_0\}) = \mathbb{C}$.

For arbitrary $\mathcal{A}$, the restriction of a pure state to a subalgebra is a pure state, and the restriction of a faithful state to a subalgebra is a faithful state. By the first paragraph, every normal $a \in \mathcal{A}$ has $\sigma(a) = \{\lambda\}$ for some $\lambda \in \mathbb{C}$. Then $\mathcal{A} = \mathbb{C}$ by Exercise 11.1.9. $\square$

**13.2.6. Irreducibility and Transitivity.** The next two results to consider, Lemma 13.2.45 and Theorem 13.2.46, are due to Kadison. The arguments we use are almost verbatim those in [**KR97a**, Lemma 5.4.2 and Theorem 5.4.3]; but we provide a simplified version that suits out needs.

***13.2.45. Lemma.***

*Let $\xi_1, \dots, \xi_n \in \mathcal{H}$ be orthonormal, and $\xi_1', \dots, \xi_n' \in \mathcal{H}$ with $\|\xi_j'\| < \delta$ for all $j$. Assume that there exists $C \in \mathcal{B}(\mathcal{H})$, selfadjoint, with $C\xi_j = \xi_j'$ for all $j$. Then there exists $B \in \mathcal{B}(\mathcal{H})$, selfadjoint, with $\|B\| \leq \sqrt{2n}\,\delta$ and $B\xi_j = \xi_j'$ for all $j$.*

*Proof.* Let $P \in \mathcal{B}(\mathcal{H})$ be the orthogonal projection onto $\overline{\operatorname{span}}\{\xi_1, \dots, \xi_n\}$. Define $X \in \mathcal{B}(\mathcal{H})$ by $X\eta = \sum_{j=1}^n \langle P\eta, \xi_j\rangle \xi_j'$. By construction, $XP = X$, so $(I - P)X^* = 0$. Also,

$$PX\eta = \sum_{j=1}^n \langle P\eta, \xi_j\rangle\, PC\xi_j = PC\,\sum_{j=1}^n \langle P\eta, \xi_j\rangle\, \xi_j = PCP\eta, \qquad \eta \in \mathcal{H}.$$

So $PX = PCP$; in particular, $PX = (PX)^* = X^*P$. Let $B = X + X^*(I - P)$. Then $B^* = X^* + X - (X^*P)^* = B$. And $B\xi_j = X\xi_j = \xi_j'$. Finally,

$$\begin{aligned}
\|X\eta\| &= \left\| \sum_{j=1}^n \langle P\eta, \xi_j\rangle \xi_j' \right\| \leq \delta \sum_{j=1}^n |\langle P\eta, \xi_j\rangle| \\
&\leq \delta \left( \sum_{j=1}^n |\langle P\eta, \xi_j\rangle|^2 \right)^{1/2} \left( \sum_{j=1}^n 1 \right)^{1/2} \\
&= \delta\,\sqrt{n}\,\|P\eta\| \leq \delta\,\sqrt{n}\,\|\eta\|,
\end{aligned}$$

so

$$\begin{aligned}\|B\|^2 = \|BB^*\| &= \|(X + X^*(I-P))(X^* + (I-P)X)\| \\ &= \|XX^* + X^*(I-P)X\| \leq 2\|X\|^2 \leq 2n\,\delta^2. \qquad \square\end{aligned}$$

Next is Kadison's remarkable Transitivity Theorem. It can be phrased by saying that topological irreducibility implies algebraic irreducibility. We present a slightly less general version than the original. For information about mapping one vector to another via a selfadjoint operator we refer to Exercise 10.3.9.

**13.2.46. Theorem. (Kadison's Transitivity Theorem)**

*Let $\mathcal{A}$ be a $C^*$-algebra and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ an irreducible representation, $\xi_1, \ldots, \xi_n \in \mathcal{H}$ linearly independent, and $\xi_1', \ldots, \xi_n' \in \mathcal{H}$. Then there exists $a \in \mathcal{A}$ such that $\pi(a)\xi_j = \xi_j'$. If the assignment $\xi_j \longmapsto \xi_j'$ can be achieved with a selfadjoint operator, then $a \in \mathcal{A}$ can be taken to be selfadjoint.*

*Proof.* Assume first that $\xi_1, \ldots, \xi_n$ is orthonormal. Let $T_0 \in \mathcal{B}(\mathcal{H})$ (selfadjoint, if possible) such that $T_0\xi_j = \xi_j'$, $j = 1, \ldots, n$. By Proposition 12.3.18 there exists $a_0 \in \mathcal{A}$ (selfadjoint, if $T_0$ is) with $\|(T_0 - \pi(a_0))\xi_j\| < (2\sqrt{2n})^{-1}$. Note that $(T_0 - \pi(a_0))\xi_j = \xi_j' - \pi(a_0)\xi_j$. Now apply Lemma 13.2.45 to $\xi_1, \ldots, \xi_n$, orthonormal, and $\xi_1' - \pi(a_0)\xi_1, \ldots, \xi_n' - \pi(a_0)\xi_n$, to obtain $T_1 \in \mathcal{B}(\mathcal{H})$ with $T_1\xi_j = \xi_j' - \pi(a_0)\xi_j$ and $\|T_1\| \leq \sqrt{2n}\,(2\sqrt{2n})^{-1} = 1/2$. If $T_0$ is selfadjoint, then $T_0 - \pi(a_0)$ is selfadjoint and so the lemma produces $T_1$ selfadjoint; when $T_0$ is not selfadjoint, we apply the lemma to its real and imaginary part, with a smaller $\delta$.

Now we work inductively. Assume that we have constructed $T_k \in \mathcal{B}(\mathcal{H})$, selfadjoint if $T_0$ is, with $\|T_k\| < 2^{-k}$ and

$$T_k\xi_j = \xi_j' - \pi(a_0)\xi_j - \pi(a_1)\xi_j - \cdots - \pi(a_{k-1})\xi_j, \qquad j = 1, \ldots, n$$

with $a_0, \ldots, a_{k-1} \in \mathcal{A}$ (selfadjoint if $T_0$ is) and $\|a_h\| < 2^{-h}$, $h = 1, \ldots, k-1$. By Proposition 12.3.18, there exists $a_k \in \mathcal{A}$, selfadjoint if $T_k$ is, with $\|a_k\| < 2^{-k}$, and $\|(T_k - \pi(a_k))\xi_j\| < (2^{k+1}\sqrt{2n})^{-1}$. This last inequality can be written as $\|\xi_j' - \sum_{h=0}^k \pi(a_h)\xi_j\| < (2^{k+1}\sqrt{2n})^{-1}$. Now Lemma 13.2.45, applied to $\xi_1, \ldots, \xi_n$ and $\left\{\xi_j' - \sum_{h=0}^k \pi(a_h)\xi_j\right\}_j$ gives us $T_{k+1} \in \mathcal{B}(\mathcal{H})$, with $\|T_{k+1}\| < \sqrt{2n}\,(2^{k+1}\sqrt{2n})^{-1} = 2^{-k-1}$ and

$$T_{k+1}\xi_j = \xi_j' - \sum_{h=0}^k \pi(a_h)\xi_j.$$

The induction is complete. Now let $a = \sum_{h=0}^{\infty} a_h \in \mathcal{A}$ (the series converges because $\|a_h\| < 2^{-h}$ for all $h \geq 1$). We have

$$\pi(a)\xi_j = \lim_{k\to\infty} \sum_{h=0}^{k} \pi(a_h)\xi_j = \xi_j' - \lim_{k\to\infty} T_{k+1}\xi_j = \xi_j', \quad j = 1, \dots, n.$$

If $T_0$ was selfadjoint, then each $a_k$ was chosen selfadjoint, and so $a = a^*$.

For the arbitrary case, where $\xi_1, \dots, \xi_n$ are linearly independent but not necessarily orthonormal, we take $\eta_1, \dots, \eta_n$ to be an orthonormal basis of $\overline{\text{span}}\{\xi_1, \dots, \xi_n\}$. Let $T_0$ be the finite-rank operator with $T_0\xi_j = \xi_j'$, and let $\eta_j' = T_0\eta_j$, $j = 1, \dots, n$. By the above, there exists $a \in \mathcal{A}$ (selfadjoint, if $T_0$ is) with $\pi(a)\eta_j = \eta_j'$ for all $j = 1, \dots, n$. Hence $\pi(a) = T_0$ on $\overline{\text{span}}\{\eta_1, \dots, \eta_n\} = \overline{\text{span}}\{\xi_1, \dots, \xi_n\}$; in particular, $\pi(a)\xi_j = T_0\xi_j = \xi_j'$. ☐

***13.2.47. Corollary.***

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be an irreducible $C^*$-algebra. If $\mathcal{A} \cap \mathcal{K}(\mathcal{H}) \neq \{0\}$, then $\mathcal{K}(\mathcal{H}) \subset \mathcal{A}$.*

*Proof.* We have that $\mathcal{B} = \mathcal{A} \cap \mathcal{K}(\mathcal{H}) \neq \{0\}$. This an ideal of $\mathcal{A}$ consisting entirely of compact operators. Taking a nonzero positive operator and using the Spectral Theorem (10.6.12), we can get a minimal projection $P \in \mathcal{B}$, so $P\mathcal{B}P = \mathbb{C}P$. This projection is necessarily rank-one. Indeed, by Kadison's Transitivity (Theorem 13.2.46), given $\xi, \eta \in PH$ there exists $B \in \mathcal{A}$ with $B\xi = \eta$. Note that $PBP$ is compact and $PBP \in \mathcal{A}$, so $PBP \in \mathcal{B}$. Then $PBP = \lambda P$ for some $\lambda \in \mathbb{C}$, and we have

$$\eta = P\eta = PB\xi = PBP\xi = \lambda P\xi = \lambda\xi,$$

implying that $P\mathcal{H}$ is one-dimensional. We then have $P = \xi\xi^*$ for some unit vector $\xi \in \mathcal{H}$. For arbitrary $\eta, \nu \in \mathcal{H}$, again by Kadison's Transitivity there exist $A, B \in \mathcal{A}$ with $A\xi = \eta$, $B\xi = \nu$. Then

$$\mathcal{B} \ni APB^* = A(\xi\xi^*)B^* = (A\xi)(B\xi)^* = \eta\nu^*.$$

So every rank-one operator is in $\mathcal{B}$, and we deduce that $\mathcal{K}(\mathcal{H}) \subset \mathcal{B}$, implying $\mathcal{A} \cap \mathcal{K}(\mathcal{H}) = \mathcal{K}(\mathcal{H})$ and thus $\mathcal{K}(\mathcal{H}) \subset \mathcal{A}$. ☐

The next proposition is due to Akemann–Anderson–Pedersen [AAP86].

***13.2.48. Proposition.*** ***(Excision)***

*Let $\varphi$ be a pure state on $\mathcal{A}$. Then there exists a net $\{e_j\} \subset \mathcal{A}$ with $0 \leq e_j \leq I$ and $\varphi(e_j) = 1$ for all $j$, and such that $\lim_j \|e_j a e_j - \varphi(a)e_j^2\| = 0$ for all $a \in \mathcal{A}$.*

*Proof.* When $\mathcal{A} = \mathbb{C}$ the result is trivial. Otherwise, we may assume that $\varphi$ is not faithful by Lemma 13.2.44. If $\mathcal{A}$ is not unital, we may replace it with $\tilde{\mathcal{A}}$; for $\varphi$ extends uniquely to $\tilde{\mathcal{A}}$ by Proposition 11.5.6 and this extension is still pure (Exercise 13.2.21) and not faithful. So we assume without loss of generality that $\mathcal{A}$ is unital.

Let $N_\varphi = \{a \in \mathcal{A} : \varphi(a^*a) = 0\}$, which is not trivial since $\varphi$ is not faithful. Because of the Cauchy–Schwarz inequality, $N_\varphi$ is a closed left ideal. Suppose that $N_\varphi + N_\varphi^* = \ker \varphi$. Since $N_\varphi \cap N_\varphi^*$ is a non-trivial C*-algebra (if $a \in N_\varphi$ is nonzero, then $|a| \in N_\varphi \cap N_\varphi^*$), it has an approximate unit $\{d_j\}$ by Theorem 11.4.4. Put $e_j = I_\mathcal{A} - d_j$; then $\varphi(e_j) = 1 - \varphi(d_j) = 1$ and, for $x \in N_\varphi$, $xe_j \to 0$, $e_j x^* \to 0$. We have, for any $a \in \mathcal{A}$, that $a - \varphi(a)I_\mathcal{A} \in \ker \varphi = N_\varphi + N_\varphi^*$; thus,

$$\|e_j a e_j - \varphi(a)e_j^2\| = \|e_j(a - \varphi(a)I_\mathcal{A})e_j\| \to 0.$$

Thus it remains to prove that $\ker \varphi = N_\varphi + N_\varphi^*$. By Cauchy–Schwarz, we always have $N_\varphi + N_\varphi^* \subset \ker \varphi$. If $a \in \ker \varphi$, we consider the (irreducible) GNS representation $(\pi, \mathcal{H}_\pi, \Omega)$ for $\varphi$. The elements $\Omega$ and $\pi(a)\Omega$ are orthogonal, since

$$\langle \pi(a)\Omega, \Omega\rangle_{\mathcal{H}_\pi} = \varphi(a) = 0.$$

By Kadison's Transitivity there exists $b \in \mathcal{A}$ , selfadjoint, such that $\pi(b)\Omega = \Omega$ and $\pi(b)\pi(a)\Omega = 0$ (we can guarantee the selfadjoint condition because the orthogonal projection onto $\mathbb{C}\Omega$ maps $\Omega \longmapsto \Omega$ and $\pi(a)\Omega \longmapsto 0$). Then $ba \in N_\varphi$ and $I_\mathcal{A} - b \in N_\varphi^*$, since

$$\varphi((ba)^*ba) = \varphi(a^*b^*ba) = \|\pi(b)\pi(a)\Omega\|^2 = 0,$$

and

$$\begin{aligned}\varphi((I_\mathcal{A} - b)aa^*(I_\mathcal{A} - b)) &\leq \|a\|^2\, \varphi((I_\mathcal{A} - b)(I_\mathcal{A} - b)) \\ &= \|a\|^2\, \|\pi(I_\mathcal{A} - b)\Omega\|^2 = \|a\|^2\, \|\Omega - \pi(b)\Omega\|^2 = 0.\end{aligned}$$

Thus $a = ba + (I_\mathcal{A} - b)a \in N_\varphi + N_\varphi^*$ and hence $\ker \varphi = N_\varphi + N_\varphi^*$. ☐

The proof of Glimm's Lemma below is taken mostly from [BO08, Lemma 1.4.11], who attribute the argument to Akitaka Kishimoto. A very different argument appears in [Dav96, Lemma II.5.1].

**13.2.49. Theorem. (Glimm's Lemma)**

*Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a separable $C^*$-algebra that contains no nonzero compact operators. Let $\varphi \in S(\mathcal{A})$. Then there exist orthonormal vectors $\{\xi_n\}$ such that*

$$\varphi(A) = \lim_n \langle A\xi_n, \xi_n\rangle, \qquad A \in \mathcal{A}.$$

*Proof.* Let $\{A_n : n \in \mathbb{N}\}$ be a dense subset of $\mathcal{A}$. We will proceed by induction: given $\{\xi_1, \ldots, \xi_n\}$ orthonormal, we will produce a unit vector $\xi_{n+1}$, orthogonal to $\xi_1, \ldots, \xi_n$, and such that $|\varphi(A_j) - \langle A_j\xi_{n+1}, \xi_{n+1}\rangle| < \frac{1}{n+1}$ for $j = 1, \ldots, n+1$.

Let $K_0 = \overline{\operatorname{span}}\{\xi_1, \ldots, \xi_n\}$. Put

$$\varepsilon = \frac{1}{(n+1)(2 + 4\max\{\|A_1\|, \ldots, \|A_{n+1}\|\})}.$$

The state space $S(\mathcal{A})$ is compact in the weak$^*$-topology (i.e., pointwise convergence). By Krein–Milman (Theorem 7.5.11) and Proposition 13.2.41, there exists a convex combination of pure states $\psi = \sum_{k=1}^m \lambda_k\psi_k$ with $|\varphi(A_j) - \psi(A_j)| < \varepsilon$, $j = 1, \ldots, n+1$. By Excision (Proposition 13.2.48) there exist $E_1, \ldots, E_m \in \mathcal{A}^+$, with $\|E_k\| = 1$, $\psi_k(E_k) = 1$, and $\|E_k(A_j - \psi_k(A_j))E_k\| < \varepsilon$.

Write $Q = I - P_{K_0}$, and let $\pi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ be the quotient map. As $P_{K_0} \in \mathcal{K}(\mathcal{H})$, $\pi(Q) = \pi(I_\mathcal{H})$. And as $\mathcal{A} \cap \mathcal{K}(\mathcal{H}) = \{0\}$, we get that $\pi$ is isometric on $\mathcal{A}$. Then

$$1 \geq \|QE_1Q\| \geq \|\pi(QE_1Q)\| = \|\pi(E_1)\| = \|E_1\| \geq |\psi_1(E_1)| = 1.$$

**Claim:** There exists a unit vector $\eta_1 \in K_0^\perp$ such that $\|E_1\eta_1 - \eta_1\| < \varepsilon$.

Assuming the claim, let

$$K_1 = \overline{\operatorname{span}}\{K_0, A_1\eta_1, \ldots, A_{n+1}\eta_1, A_1^*\eta_1, \ldots, A_{n+1}^*\eta_1\}.$$

Repeating the above argument with $K_1$ instead of $K_0$, from the claim we obtain $\eta_2 \in K_1^\perp$ with $\|E_2\eta_2 - \eta_2\| < \varepsilon$. And then, iterating this construction, we end up obtaining $\eta_1, \ldots, \eta_m$, orthonormal, with $\|E_k\eta_k - \eta_k\| < \varepsilon$, $k = 1, \ldots, m$. Let $\xi_{n+1} = \sum_{k=1}^m \lambda_k^{1/2}\eta_k$. Then $\|\xi_{n+1}\| = 1$ and, for $j = 1, \ldots, n+1$,

$$\begin{aligned}
|\varphi(A_j) - \langle A_j\xi_{n+1}, \xi_{n+1}\rangle| &= \left|\varphi(A_j) - \sum_{k=1}^m \lambda_k\langle A_j\eta_k, \eta_k\rangle\right| \\
&\leq |\varphi(A_j) - \psi(A_j)| + \sum_{k=1}^m \lambda_k|(\psi_k(A_j) - \langle A_j\eta_k, \eta_k\rangle| \\
&\leq \varepsilon + \sum_{k=1}^m \lambda_k|\langle(\psi_k(A_j) - A_j)\eta_k, \eta_k\rangle|
\end{aligned}$$

$$\leq \varepsilon + \sum_{k=1}^{m} \lambda_k \|E_k(A_j - \psi_k(A_j)I)E_k\|$$
$$+ \varepsilon\|(A_j - \psi_k(A_j))E_k\| + \varepsilon\|A_j - \psi_k(A_j)I\|$$
$$\leq (2 + 4\|A_j\|)\,\varepsilon < \frac{1}{n+1}.$$

It remains to prove the claim. Fix $\delta > 0$, such that $\frac{\delta^{1/2}+1-(1-\delta)^{1/2}}{(1-\delta)^{1/2}} < \varepsilon$. Since $\|QE_1Q\| = 1$ and $QE_1Q \geq 0$, we have $1 \in \sigma(QE_1Q)$. So we can find a unit vector $\zeta \in \mathcal{H}$ with $\|QE_1Q\zeta - \zeta\| < 1 - (1-\delta)^{1/2}$ (since for a positive operator the spectrum agrees with the approximate point spectrum). Then

$$1 - \|QE_1Q\zeta\| = \|\zeta\| - \|QE_1Q\zeta\| \leq \|\zeta - QE_1Q\zeta\| \leq 1 - (1-\delta)^{1/2}.$$

We deduce that $(1-\delta)^{1/2} \leq \|QE_1Q\zeta\|$. So

$$1 - \delta \leq \|QE_1Q\zeta\|^2 \leq \|E_1Q\zeta\|^2 \leq 1. \tag{13.10}$$

Since $\|QE_1Q\zeta\|^2 + \|Q^\perp E_1Q\zeta\|^2 = \|QE_1Q\zeta + Q^\perp E_1Q\zeta\|^2 = \|E_1Q\zeta\|^2$, we get from (13.10) that

$$\|Q^\perp E_1Q\zeta\|^2 = \|E_1Q\zeta\|^2 - \|QE_1Q\zeta\|^2 \leq 1 - (1-\delta) = \delta.$$

Thus

$$\|E_1Q\zeta - Q\zeta\| \leq \|Q^\perp E_1Q\zeta\| + \|QE_1Q\zeta - Q\zeta\|$$
$$\leq \delta^{1/2} + \|Q(QE_1Q\zeta - \zeta)\| \leq \delta^{1/2} + \|QE_1Q\zeta - \zeta\|$$
$$\leq \delta^{1/2} + 1 - (1-\delta)^{1/2}.$$

Now let $\eta_1 = Q\zeta/\|Q\zeta\|$. Note that $\|Q^\perp\zeta\| = \|Q^\perp(\zeta - QE_1Q\zeta)\| < 1 - (1-\delta)^{1/2}$. So

$$\|Q\zeta\| = \|\zeta - Q^\perp\zeta\| \geq 1 - \|Q^\perp\zeta\| \geq 1 - (1 - (1-\delta)^{1/2}) = (1-\delta)^{1/2}.$$

Then

$$\|E_1\eta_1 - \eta_1\| = \frac{\|E_1Q\zeta - Q\zeta\|}{\|Q\zeta\|} \leq \frac{\delta^{1/2} + 1 - (1-\delta)^{1/2}}{(1-\delta)^{1/2}} < \varepsilon. \qquad \square$$

**13.2.7. Completely Positive Maps on $M_n(\mathbb{C})$.** Possibly the most effective way of determining whether a linear map $\phi : M_n(\mathbb{C}) \to \mathcal{B}(\mathcal{H})$ is completely positive is the following result from Choi. We identify $M_n(M_m(\mathbb{C}))$ with the span of $E_{kj}^{(m)} \otimes E_{rs}^{(n)}$.

***13.2.50. Theorem. (Choi's Criterion)***

*Let $\phi : M_n(\mathbb{C}) \to \mathcal{B}(\mathcal{H})$ be a linear map. The following statements are equivalent:*

*(1)* $\phi$ *is completely positive;*
*(2)* $\phi$ *is n-positive;*
*(3)* $\phi^{(n)}\left(\sum_{k,j=1}^n E_{kj}\otimes E_{kj}\right)\geq 0.$

*Proof.* (1) $\implies$ (2) Trivial.

(2) $\implies$ (3) We will verify that $A=\sum_{k,j=1}^n E_{kj}\otimes E_{kj}$ is positive, and then the result follows from the $n$-positivity of $\phi$. The matrix $A$ is clearly selfadjoint. We also have

$$A^2=\sum_{s,t=1}^n\sum_{k,j=1}^n E_{kj}E_{ts}\otimes E_{kj}E_{ts}=\sum_{s=1}^n\sum_{k,j=1}^n E_{ks}\otimes E_{ks}=n\,A.$$

This implies $A=(A/\sqrt{n})^2$, so $A$ is positive.

(3) $\implies$ (1) We will apply Lemma 13.2.10 for each $m\in\mathbb{N}$. Fix matrices $A_1,\dots,A_m\in M_n(\mathbb{C})$, $\xi_1,\dots,\xi_m\in\mathbb{C}^n$. For each $A_k$, write $A_k=\sum_{s,t=1}^m a_{kst}E_{st}$. Let $\eta_1,\dots,\eta_n$ denote the canonical basis of $\mathbb{C}^n$. Then

$$\begin{aligned}
\sum_{k,j=1}^m\langle\phi(A_j^*A_k)\xi_k,\xi_j\rangle &= \sum_{k,j=1}^m\sum_{s,t=1}^n\sum_{x,y=1}^n\overline{a_{jst}}a_{kxy}\langle\phi(E_{ts}E_{xy})\xi_k,\xi_j\rangle\\
&=\sum_{k,j=1}^m\sum_{s,t=1}^n\sum_{y=1}^n\overline{a_{jst}}a_{ksy}\langle\phi(E_{ty})\xi_k,\xi_j\rangle\\
&=\sum_{s=1}^n\left[\sum_{t,y=1}^n\left\langle\phi(E_{ty})\left(\sum_{k=1}^m a_{ksy}\xi_k\right),\left(\sum_{j=1}^m a_{jst}\xi_j\right)\right\rangle\right]\\
&=\sum_{s=1}^n\left\langle\phi^{(n)}\left(\sum_{t,y=1}^n E_{ty}\otimes E_{ty}\right)\tilde{\xi}_s,\tilde{\xi}_s\right\rangle\geq 0,
\end{aligned}$$

where $\tilde{\xi}_s=\sum_{t=1}^n\left(\sum_{k=1}^m a_{kst}\xi_k\right)\otimes\eta_t$. Then $\phi$ is $m$-positive for each $m\in\mathbb{N}$. $\square$

For better visualization, note that the Choi Matrix is

$$\phi^{(n)}\left(\sum_{k,j=1}^n E_{kj}\otimes E_{kj}\right)=\begin{bmatrix}\phi(E_{11}) & \phi(E_{12}) & \cdots & \phi(E_{1n})\\ \phi(E_{21}) & \phi(E_{22}) & \cdots & \phi(E_{2n})\\ \vdots & & & \\ \phi(E_{n1}) & \phi(E_{n2}) & \cdots & \phi(E_{nn})\end{bmatrix}. \tag{13.11}$$

It is remarkable that complete positivity for a map can be tested on a single matrix. In the case where $\mathcal{H}$ is finite-dimensional, we can use this fact to obtain the following characterization:

***13.2.51. Corollary. (Kraus Decomposition)***

*Let $\phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ be completely positive. Then there exist matrices $V_1, \ldots, V_{nm} \in B(\mathbb{C}^n, \mathbb{C}^m)$ with*

$$\phi(X) = \sum_{\ell=1}^{nm} V_\ell X V_\ell^*, \qquad X \in M_n(\mathbb{C}).$$

*Proof.* Write $X = \sum_{k,j=1}^{n} x_{kj}\, E_{kj} \in M_n(\mathbb{C})$. Let

$$Y = \phi^{(n)}\left(\sum_{k,j=1}^{n} E_{kj}^{(n)} \otimes E_{kj}^{(n)}\right) \in M_n(M_m(\mathbb{C})).$$

By Theorem 13.2.50 $Y$ is positive, so there exists $R \in M_n(M_m(\mathbb{C}))$ with $Y = R^*R$. Then, seeing $Y, R$ as $n \times n$ block matrices with $m \times m$ blocks,

$$\begin{aligned}
\phi(X) &= \sum_{k,j=1}^{n} x_{kj} Y_{kj} = \sum_{k,j=1}^{n} x_{kj} \sum_{\ell=1}^{n} R_{\ell k}^* R_{\ell j} = \sum_{\ell=1}^{n} \sum_{k,j=1}^{n} x_{kj} R_{\ell k}^* R_{\ell j} \\
&= \sum_{\ell=1}^{n} \begin{bmatrix} R_{\ell 1}^* & \cdots & R_{\ell n}^* \end{bmatrix} \begin{bmatrix} x_{11} I_m & \cdots & x_{1n} I_m \\ \vdots & \ddots & \vdots \\ x_{n1} I_m & \cdots & x_{nn} I_m \end{bmatrix} \begin{bmatrix} R_{\ell 1} \\ \vdots \\ R_{\ell n} \end{bmatrix} \\
&= \sum_{\ell=1}^{n} S_\ell (I_m \otimes X) S_\ell^*
\end{aligned}$$

where $S_\ell = \begin{bmatrix} R_{\ell 1}^* & R_{\ell 2}^* & \cdots & R_{\ell n}^* \end{bmatrix}$ (i.e., $S_{\ell k} = R_{\ell k}^*$). Let $W \in M_{nm}(\mathbb{C})$ be the unitary that implements the flip: $W^*(I_m \otimes X)W = X \otimes I_m$, independently of $X$ (it is just a reorder of the canonical basis, so it is implemented by a unitary and it does not depend on $X$). Let $T_\ell = S_\ell W$. Note that each $T_\ell$ is an $m \times nm$ matrix. Instead of thinking of it as a block $1 \times n$ with $m \times m$ blocks, we can write it as a block $1 \times m$ matrix with $m \times n$ blocks, i.e. $T_\ell = \begin{bmatrix} V_{\ell 1} & V_{\ell 2} & \cdots & V_{\ell m} \end{bmatrix}$, with each $V_{\ell k}$ an $m \times n$ matrix. Then, multiplying $T_\ell (X \otimes I_m) T_\ell^*$ as $1 \times m$ by $m \times m$ by $m \times 1$ block matrices,

$$\phi(X) = \sum_{\ell=1}^{n} T_\ell (X \otimes I_m) T_\ell^* = \sum_{\ell=1}^{n} \sum_{k,j=1}^{m} V_{\ell k} (X \otimes I_m)_{kj} V_{\ell j}^* = \sum_{\ell=1}^{n} \sum_{k=1}^{m} V_{\ell k} X V_{\ell k}^*. \square$$

*13.2.52. Remark*. The decomposition from Corollary 13.2.51 is not unique. For a simple example, consider $\phi : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ given by

$$\phi\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}.$$

This map is obviously ucp, since

$$\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

We also have

$$\begin{aligned} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & -1/\sqrt{2} \end{bmatrix} \\ &\quad + \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix}. \end{aligned}$$

None of these two decompositions have the predicted number of 4 terms. But of course this can be achieved by completing with zeroes. Here is a third decomposition, this time with four terms.

$$\begin{aligned} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}. \end{aligned}$$

**13.2.8. Completely positive maps into $M_n(\mathbb{C})$.** One of the reasons why completely positive maps are the right generalization of linear functionals is a "Hahn–Banach" result of Arveson: the celebrated and useful **Arveson's Extension Theorem** (13.2.59 below). The proof of the theorem will require two technical results: the first one is a technique to extend completely positive maps with finite-dimensional range, while the second one is to develop a way to find accumulation points of those maps.

The key to the first step is a certain duality between completely positive maps and linear functionals. Given a linear map $\phi : \mathcal{S} \to M_n(\mathbb{C})$ we associate to it a linear functional $f_\phi : M_n(\mathcal{S}) \to \mathbb{C}$ by

$$f_\phi(X) = \frac{1}{n} \left\langle \phi^{(n)}(X)e, e \right\rangle, \tag{13.12}$$

where $e = \xi_1 + \cdots + \xi_n$ and $\{\xi_1, \ldots, \xi_n\}$ is the canonical basis of $\mathbb{C}^n$. It is easy to see that

$$f_\phi(X) = \frac{1}{n} \sum_{k,j=1}^{n} \phi(x_{kj})_{kj}.$$

Conversely, given a functional $f : M_n(\mathcal{S}) \to \mathbb{C}$, we obtain a map $\phi_f : \mathcal{S} \to M_n(\mathbb{C})$ by

$$\phi_f(s) = n \sum_{k,j=1}^{n} f(s \otimes E_{k,j})\, E_{k,j} \tag{13.13}$$

These two constructions are inverses of each other:

**13.2.53. Proposition.**

*Let $\phi : \mathcal{S} \to M_n(\mathbb{C})$ be a linear map. Then $\phi_{f_\phi} = \phi$. Conversely, given a functional $f : M_n(\mathcal{S}) \to \mathbb{C}$, we have $f_{\phi_f} = f$.*

*Proof.* For $s \in \mathcal{S}$,

$$\phi_{f_\phi}(s) = n \sum_{k,j=1}^{n} f_\phi(s \otimes E_{kj})\, E_{kj} = \sum_{k,j=1}^{n} \phi(s)_{kj}\, E_{kj} = \phi(s).$$

And, for $X \in M_n(\mathcal{S})$,

$$f_{\phi_f}(X) = \frac{1}{n} \sum_{k,j=1}^{n} \phi_f(x_{kj})_{kj} = \sum_{k,j=1}^{n} f(x_{kj} \otimes E_{kj}) = f(X). \qquad \square$$

The importance of this duality stems from the fact that it shows that the set of completely positive maps from an operator system into the matrices,

$$CP(\mathcal{S}) = \{\phi : \mathcal{S} \to M_n(\mathbb{C}) : \ n \in \mathbb{N}, \ \phi \ \text{ completely positive}\},$$

is encoded by the set of all the positive functionals on $M_n(\mathcal{S})$, for all $n$. In particular,

**13.2.54. Theorem.**

*Let $\phi : \mathcal{S} \to M_n(\mathbb{C})$ be a linear map. The following statements are equivalent:*

*(i) $\phi$ is completely positive;*

*(ii) $\phi$ is $n$-positive;*

*(iii) $f_\phi$ is positive.*

*Proof.*

(i) $\implies$ (ii) Trivial.

(ii) $\implies$ (iii) This follows directly from the definition in (13.12).

(iii) $\implies$ (i) By Theorem 13.2.11 the positive linear functional $f_\phi$ can be extended to positive linear functional of $M_n(C^*(\mathcal{S}))$. Being positive, it is completely positive by Proposition 13.2.12. Now fix $X \in M_m(C^*(\mathcal{S}))$, positive. Then $X = Y^*Y$ for some $Y \in M_m(C^*(\mathcal{S}))$. Using that $\phi = \phi_{f_\phi}$ by Proposition 13.2.53,

$$\phi^{(m)}(X) = \phi^{(m)}(Y^*Y) = \sum_{k,j=1}^{m} \phi((Y^*Y)_{kj}) \otimes E_{kj}^{m}$$

$$
\begin{aligned}
&= \sum_{k,j=1}^{m} \sum_{\ell=1}^{m} \phi(Y_{\ell k}^* Y_{\ell j}) \otimes E_{kj}^m \\
&= n \sum_{\ell=1}^{m} \sum_{k,j=1}^{m} \sum_{t,s=1}^{n} f_\phi(Y_{\ell k}^* Y_{\ell j} \otimes E_{ts}^n) E_{ts}^n \otimes E_{kj}^m \\
&= n \sum_{\ell=1}^{m} \sum_{k,j=1}^{m} f_\phi^{(n)} \left( \sum_{t,s=1}^{n} (Y_{\ell k}^* Y_{\ell j} \otimes E_{ts}^n) \otimes E_{ts}^n \right) \otimes E_{kj}^m \\
&= n \sum_{\ell=1}^{m} \sum_{k,j=1}^{m} f_\phi^{(n)} \left(B_{\ell k}^* B_{\ell j}\right) \otimes E_{kj}^m \\
&= n \sum_{\ell=1}^{m} f_\phi^{(nm)} \left( \left( \sum_{k=1}^{m} B_{\ell k} \otimes E_{1k}^m \right)^* \left( \sum_{k=1}^{m} B_{\ell k} \otimes E_{1k}^m \right) \right) \geq 0,
\end{aligned}
$$

where $B_{\ell k} = \sum_{t=1}^{n} Y_{\ell k} \otimes E_{1t}^n \otimes E_{1t}^n$. □

*13.2.55. Remark.* Let $\phi : \mathcal{S} \to M_n(\mathbb{C})$ be completely positive. For $x \in \mathcal{S}$, using that $\|\phi\|_{\mathrm{cb}} = \|\phi\|$ (Proposition 13.2.24),

$$|f_\phi(x)| = \frac{1}{n} |\langle \phi^{(n)}(x)e, e\rangle| \leq \frac{1}{n} \|\phi^{(n)}\| \, \|x\| \, \|e\|^2 = \|\phi\| \, \|x\|.$$

So $\|f_\phi\| \leq \|\phi\|$. But for a positive $f : M_n(\mathcal{S}) \to \mathbb{C}$ there is no such nice relation between $\|\phi_f\|$ and $\|f\|$; indeed, consider $f : M_n(\mathbb{C}) \to \mathbb{C}$ given by $f(a) = \frac{1}{n} \sum_{k,j=1}^{n} a_{kj}$. Then

$$\phi_f(1) = n \sum_{k,j=1}^{n} f(1 \otimes E_{kj}) E_{kj} = \sum_{k,j=1}^{n} E_{k,j},$$

so $\|f\| = f(I_n) = 1$ and $\|\phi_f\| = \|\phi_f(1)\| = n$.

Theorem 13.2.54 allows us to prove the following extension result, which plays a key role in the proof Arveson's Extension Theorem 13.2.59.

***13.2.56. Corollary.***

*Let $\phi : \mathcal{S} \to M_n(\mathbb{C})$ be completely positive. If $\mathcal{S} \subset \mathcal{A}$, then there exists a completely positive extension $\tilde{\phi} : \mathcal{A} \to M_n(\mathbb{C})$ of $\phi$, i.e. $\tilde{\phi}|_{\mathcal{S}} = \phi$.*

*Proof.* By Theorem 13.2.54, $f_\phi : M_n(\mathcal{S}) \to \mathbb{C}$ is positive. So we can extend it by Theorem 13.2.11 to a positive linear functional $g : M_n(\mathcal{A}) \to \mathbb{C}$. By

Theorem 13.2.54, the map $\phi_g : \mathcal{A} \to M_n(\mathbb{C})$ is completely positive. For $s \in \mathcal{S}$,

$$\phi_g(s) = n \sum_{k,j=1}^{n} g(s \otimes E_{kj})E_{kj} = n \sum_{k,j=1}^{n} f_\phi(s \otimes E_{kj})E_{kj} = \phi_{f_\phi}(s) = \phi(s)$$

by Proposition 13.2.53, so $\phi_g$ extends $\phi$. □

**13.2.9. Arveson's Extension Theorem.** We will write $CP(\mathcal{S}, \mathcal{B})$ for the set

$$CP(\mathcal{S}, \mathcal{B}) = \{\phi : \mathcal{S} \to \mathcal{B} : \ \phi \text{ is cp }\}.$$

Our first goaln is to give this convex set a topological structure. A straightforward choice would be to use the cb-norm $\|\cdot\|_{\text{cb}}$. The problem is, as with the norm on $\mathcal{B}(\mathcal{H})$, that closed balls are not compact; this is an issue for existence theorems.

So Arveson's idea, likely inspired by [Kad61], was to build a weak* topology for $CP(\mathcal{S}, \mathcal{B})$. This requires it to be a dual; which may be hard in general, but not so much when $\mathcal{B} = \mathcal{B}(\mathcal{H})$. And, as we have discussed in Remark 13.2.7, we do not lose much—if anything—by taking $\mathcal{B} = \mathcal{B}(\mathcal{H})$.

By Theorem 10.7.11 $\mathcal{B}(\mathcal{H})$ can be identified with the dual of the Banach space $\mathcal{T}(H)$, the trace-class operators, via the duality

$$\hat{T}(S) = \text{Tr}(TS), \quad T \in \mathcal{B}(\mathcal{H}), \ S \in \mathcal{T}(H).$$

We called the weak*-topology on $\mathcal{B}(\mathcal{H})$, when it is considered as the dual $\mathcal{T}(H)$, the **ultraweak topology**, or the **$\sigma$-weak topology**. So $CP(\mathcal{S}, \mathcal{B}(\mathcal{H})) = CP(\mathcal{S}, \mathcal{T}(H)^*)$. In order to make Banach space considerations, we need to embed $CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$—which is a cone, but not a vector space—into a Banach space. Let

$$B(\mathcal{S}, \mathcal{B}(\mathcal{H})) = \{\phi : \mathcal{S} \to \mathcal{B}(\mathcal{H}) : \ \phi \text{ is linear and bounded }\}.$$

**13.2.57. Definition.**

*The **BW-topology** on $B(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ is the topology given by pointwise weak*-convergence. That is, $\varphi_j \to \varphi$ means that $\text{Tr}(T\varphi_j(s)) \to \text{Tr}(T\varphi(s))$ for all $s \in \mathcal{S}$ and $T \in \mathcal{T}(\mathcal{H})$.*

BW stands for "bounded-weak". It turns out that we have already considered this situation in Section 9.4.

***13.2.58. Proposition.***

*The unit ball of $CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$, i.e. the set of completely positive contractions $\mathcal{S} \to \mathcal{B}(\mathcal{H})$, is BW-compact.*

*Proof.* By Proposition 9.4.10, the closed unit ball of $B(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ is compact in the BW-topology. So all we need to show is that the unit ball of $CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ is BW- closed. Let $\{\phi_j\}$ be a BW-convergent net in the unit ball of $CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$, so $\phi_j \xrightarrow{BW} \phi$ for some $\phi \in B(\mathcal{S}, \mathcal{B}(\mathcal{H}))$. For unit vectors $\xi, \eta \in \mathcal{H}$ and $s \in \mathcal{S}$

$$|\langle \phi(s)\xi, \eta\rangle| = |\lim_j \langle \phi_j(s)\xi, \eta\rangle| \leq \|s\|,$$

so $\|\phi\| \leq 1$, so we need to check its complete positivity. Using Exercise 13.2.23 we have, for $a \in M_n(\mathcal{S})^+$, $\xi \in \mathcal{H}^n$,

$$\begin{aligned}\langle \phi^{(n)}(a)\xi, \xi\rangle &= \sum_{k,h=1}^{n} \langle \phi(a_{kh})\xi_h, \xi_k\rangle = \lim_j \sum_{k,h=1}^{n} \langle \phi_j(a_{kh})\xi_h, \xi_k\rangle \\ &= \lim_j \langle \phi_j^{(n)}(a)\xi, \xi\rangle \geq 0.\end{aligned}$$

Thus, $\phi$ is completely positive and so $CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ is closed. $\square$

Before stating and proving the Extension Theorem it is worth noting that positive maps often fail to have positive extensions. Consider for example the map $\phi$ from Example 13.2.9. If there was an extension $\tilde{\phi}$ to $C(\overline{\mathbb{D}})$, then by Proposition 13.2.22 we would have that $\tilde{\phi}$ is completely positive and unital, and then Theorem 13.2.19 would imply $\|\tilde{\phi}\| = 1$, contradicting the fact that $\|\phi\| = 2$.

With the material collected about extensions of completely positive maps into $M_n(\mathbb{C})$ and the BW topology, we have the necessary ingredients to prove Arveson's Extension Theorem:

***13.2.59. Theorem. (Arveson* [Arv69]*)***

*Let $\phi : \mathcal{S} \to \mathcal{B}(\mathcal{H})$ be a completely positive map. If $\mathcal{S} \subset \mathcal{A}$, then there exists a completely positive map $\tilde{\phi} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ such that $\tilde{\phi}|_{\mathcal{S}} = \phi$.*

*Proof.* Let $\{P_j\}$ be an increasing net of finite-rank projections with $P_j\xi \to \xi$ for all $\xi \in \mathcal{H}$. For each $P_j$, we have a $*$-isomorphism $\pi_j : P_j\mathcal{B}(\mathcal{H})P_j \to M_{k(j)}(\mathbb{C})$ (Exercise 13.2.24). Let $\phi_j : \mathcal{S} \to M_{k(j)}(\mathbb{C})$ be given by $\phi_j(x) = \pi_j(P_j\phi(x)P_j)$; this map is completely positive since it is a composition of completely positive maps (see Examples 13.2.6).

By Corollary 13.2.56, there exist cp maps $\tilde{\phi}_j : \mathcal{A} \to M_{k(j)}(\mathbb{C})$ with $\tilde{\phi}_j|_{\mathcal{S}} = \phi_j$. Since they are completely positive and so are the maps $\pi_j^{-1}$, by Proposition 13.2.24 we have

$$\begin{aligned}\|\pi_j^{-1} \circ \tilde{\phi}_j\| &= \|\pi_j^{-1} \circ \tilde{\phi}_j(I_{\mathcal{A}})\| = \|\pi_j^{-1} \circ \phi_j(I_{\mathcal{A}})\| \\ &= \|\pi_j^{-1} \circ \pi_j(P_j\phi(I_{\mathcal{A}})P_j)\| \leq \|\phi(I_{\mathcal{A}})\|.\end{aligned}$$

The boundedness of the net $\{\pi_j^{-1}\circ\phi_j\}$ guarantees that its BW closure is compact (Proposition 13.2.58), and so there is a completely positive accumulation point $\tilde\phi:\mathcal{A}\to\mathcal{B}(\mathcal{H})$. By replacing the net $\{\pi_j^{-1}\circ\tilde\phi_j\}$ with the corresponding subnet, we might assume that $\pi_j^{-1}\circ\tilde\phi_j\xrightarrow{BW}\tilde\phi$. So $\tilde\phi$ is a completely positive map; and, for $x\in\mathcal{S}$, $\xi\in\mathcal{H}$,

$$\begin{aligned}\langle\tilde\phi(x)\xi,\xi\rangle&=\lim_j\langle\pi_j^{-1}\circ\tilde\phi_j(x)\xi,\xi\rangle=\lim_j\langle P_j\phi(x)P_j\xi,\xi\rangle\\&=\lim_j\langle\phi(x)P_j\xi,P_j\xi\rangle=\langle\phi(x)\xi,\xi\rangle,\end{aligned}$$

where the last equality is justified by the fact that $P_j\xi\to\xi$ and the net $\{P_j\xi\}$ is bounded. As $\xi$ was arbitrary, we conclude that $\tilde\phi(x)=\phi(x)$ for all $x\in\mathcal{S}$. ☐

*13.2.60. Remark*. Because of Proposition 13.2.12 and Theorem 13.2.19, and the fact that operator systems contain the identity operator, Theorem 13.2.11 is a particular case (for $\dim\mathcal{H}=1$) of Theorem 13.2.59.

**13.2.10. Further Technical Results.** We collect here some results about completely positive maps that will be needed down the road. We will systematically see $\mathcal{A}^{**}$ as a C*-algebra as in Theorem 12.7.8.

***13.2.61. Lemma.***

*Let $\mathcal{A},\mathcal{B}$ be unital C*-algebras, $\theta:\mathcal{A}\to\mathcal{B}$ a unital $*$-homomorphism, and $\{\varphi_j\}$ and $\{\psi_j\}$ nets of contractive completely positive maps $\varphi_j:\mathcal{A}\to M_{n(j)}(\mathbb{C})$ and $\psi_j:M_{n(j)}(\mathbb{C})\to\mathcal{B}$ with $\psi_j\circ\varphi_j\to\theta$ pointwise. Then the maps $\varphi_j$ and $\psi_j$ can be replaced, after possibly shrinking the $n(j)$, with ucp maps that still satisfy the same convergence.*

*Proof.* The element $\varphi_j(I_\mathcal{A})\in M_{n(j)}(\mathbb{C})$ is positive. Being selfadjoint, its range projection is the orthogonal of its kernel projection. Let $P_j\in M_{n(j)}(\mathbb{C})$ be its range-projection. This implies $\varphi_j(I_\mathcal{A})$ is invertible in $P_jM_{n(j)}(\mathbb{C})P_j$. So there exist matrices $Z_j\in M_{n(j)}(\mathbb{C})$ such that $Z_j\varphi_j(I_\mathcal{A})=P_j$ for each $j$; since $\varphi_j(I_\mathcal{A})\geq 0$ we also have $Z_j\geq 0$. We have, for $a\geq 0$,

$$\begin{aligned}0&\leq(I_{n(j)}-P_j)\varphi_j(a)(I_{n(j)}-P_j)\leq(I_{n(j)}-P_j)\varphi_j(\|a\|\,I_\mathcal{A})(I_{n(j)}-P_j)\\&=\|a\|\,(I_{n(j)}-P_j)P_j\varphi_j(I_\mathcal{A})(I_{n(j)}-P_j)=0.\end{aligned}$$

Hence $0=\varphi_j(a)^{1/2}(I_{n(j)}-P_j)$ and we get that $\varphi_j(a)=P_j\varphi_j(a)$; as the positive elements span $\mathcal{A}$, the equality holds for all $a\in\mathcal{A}$. We define

$$\varphi_j'(a)=Z_j^{1/2}\varphi_j(a)Z_j^{1/2}.$$

Then $\varphi_j'$ is completely positive and $\varphi_j'(I_{\mathcal{A}}) = Z_j^{1/2}\varphi_j(I_{\mathcal{A}})Z_j^{1/2} = P_j$. If we replace $M_{n(j)}(\mathbb{C})$ with $P_jM_{n(j)}(\mathbb{C})P_j \simeq M_{\mathrm{Tr}(P_j)}(\mathbb{C})$, now $\varphi_j'$ is unital and we replace $n(j)$ with $\mathrm{Tr}(P_j)$. We have $\psi_j(P_j) = \psi_j(\varphi_j(I_{\mathcal{A}})) \to \theta(I_{\mathcal{A}}) = I_{\mathcal{B}}$. So, for $j$ big enough, $\psi_j(P_j)$ is invertible (by Proposition 6.2.3) and positive. If we only consider $j$ such that $\|I_{\mathcal{B}} - \psi_j(P_j)\| < \frac{1}{2}$, then $\psi_j(P_j) \geq \frac{1}{2}\, I_{\mathcal{B}}$ for all such $j$. From the equality

$$1 - t^{-1/2} = -t^{-1/2}(1 + t^{1/2})^{-1}(1 - t), \qquad\qquad t \in (0,1)$$

and functional calculus we get

$$\|I_{\mathcal{B}} - \psi_j(P_j)^{-1/2}\| \leq \sqrt{2}\,\|I_{\mathcal{B}} - \psi_j(P_j)\| \xrightarrow[j]{} 0.$$

We define

$$b_j = \psi_j(\varphi_j(I_{\mathcal{A}}))^{-1/2} \in \mathcal{B}, \qquad\qquad Y_j = \varphi_j(I_{\mathcal{A}})^{1/2} \in M_{n(j)}(\mathbb{C}),$$

and

$$\psi_j'(X) = b_j\psi_j(Y_jXY_j)b_j, \qquad\qquad X \in M_{n(j)}(\mathbb{C}).$$

Now

$$\psi_j'(I_{n(j)}) = b_j\psi_j(Y_j^2)b_j = \psi_j(\varphi_j(I_{\mathcal{A}}))^{-1/2}\psi_j(\varphi_j(I_{\mathcal{A}}))\psi_j(\varphi_j(I_{\mathcal{A}}))^{-1/2} = I_{\mathcal{B}},$$

so $\psi_j'$ is unital. Also, since $b_j \to I_{\mathcal{B}}$ and $Z_j^{1/2}Y_j = P_j$ (by taking the square root of $Z_jY_j^2 = P_j$, since they are positive and commute),

$$\begin{aligned}\lim_j \|\psi_j'(\varphi_j'(a)) - \theta(a)\| &= \lim_j \|b_j\psi_j(Y_jZ_j^{1/2}\varphi_j(a)Z_j^{1/2}Y_j)b_j - \theta(a)\| \\ &= \lim_j \|\psi_j(P_j\varphi_j(a)P_j) - \theta(a)\| \\ &= \lim_j \|\psi_j(\varphi_j(a)) - \theta(a)\| = 0. \qquad \square\end{aligned}$$

***13.2.62. Lemma.***

*Let $\mathcal{A}$ be a $C^*$-algebra and $\{\varphi_j\}$ and $\{\psi_j\}$ nets of contractive completely positive maps $\varphi_j : \mathcal{A} \to M_{n(j)}(\mathbb{C})$ and $\psi_j : M_{n(j)}(\mathbb{C}) \to \mathcal{A}^{**}$ with $\psi_j \circ \varphi_j \to \mathrm{id}_{\mathcal{A}^{**}}$ pointwise weak*. Then the maps $\varphi_j$ and $\psi_j$ can be replaced with ucp maps that still satisfy the same convergence.*

*Proof.* We get unital maps $\varphi_j'$ as in the first part of the proof of Lemma 13.2.61 (and at the same time replacing $M_{n(j)}(\mathbb{C})$ with a subalgebra if needed, which in terms of notation amounts to redefine $n(j)$). Since the maps are positive and contractive, we have $\psi_j(\varphi_j(I_{\mathcal{A}})) = \psi_j(I_{n(j)}) \leq I_{\mathcal{A}^{**}}$. Also,

$$\psi_j(\varphi_j(I_{\mathcal{A}})) \xrightarrow{\text{weak}^*} I_{\mathcal{A}^{**}}$$

(recall from Exercise 12.7.4 that $I_{\mathcal{A}}$ is identified with $I_{\mathcal{A}^{**}}$ in $\mathcal{A}^{**}$). So if we define $b_j = I_{\mathcal{A}^{**}} - \psi_j(\varphi_j(I_{\mathcal{A}})) \in \mathcal{A}^{**}$ we get positive elements that converge weak* to zero. Let $\rho_j$ be the state on $M_{n(j)}(\mathbb{C})$ that returns the $1,1$ entry (any other state would do), and define $\psi'_j : M_{n(j)}(\mathbb{C}) \to \mathcal{A}^{**}$ by

$$\psi'_j(X) = \rho_j(X)b_j + \psi_j(\varphi_j(I_{\mathcal{A}})^{1/2} X \varphi_j(I_{\mathcal{A}})^{1/2})$$

Then $\psi_j$ is unital and completely positive by construction, and (in the weak*-topology)

$$\begin{aligned}\lim_j \psi'_j(\varphi'_j(a)) &= \lim_j \rho_j(\varphi'_j(a))b_j + \psi_j\big(\varphi_j(I_{\mathcal{A}})^{1/2}\varphi'_j(a)\varphi_j(I_{\mathcal{A}})^{1/2}\big) \\ &= \lim_j \psi_j(\varphi_j(I_{\mathcal{A}})^{1/2} Z_j^{1/2} \varphi_j(a) Z_j^{1/2} \varphi_j(I_{\mathcal{A}})^{1/2}) \\ &= \lim_j \psi_j(P_j\varphi_j(a)P_j) = \lim_j \psi_j(\varphi_j(a)) = a. \qquad \square\end{aligned}$$

***13.2.63. Lemma.***

*Let $\mathcal{A}$ be a $C^*$-algebra and $\theta_1, \dots, \theta_n : \mathcal{A} \to \mathcal{A}$ completely positive and such that for each $k$ there exist $\varphi : \mathcal{A} \to M_{n(k)}(\mathbb{C})$ and $\psi : M_{n(k)}(\mathbb{C}) \to \mathcal{A}$ completely positive with $\theta_k = \psi_k \circ \varphi_k$. Then for each convex combination $\theta = \sum_k t_k\theta_k$ there exist $r \in \mathbb{N}$ and completely positive maps $\varphi : \mathcal{A} \to M_r(\mathbb{C})$, $\psi : M_r(\mathbb{C}) \to \mathcal{A}$ such that $\theta = \psi \circ \varphi$. The maps $\varphi$ and $\psi$ can be chosen contractive if the original $\varphi_k$ and $\psi_k$ are.*

*Proof.* It is enough to prove the assertion when $n = 2$. Let $r = n(1) + n(2)$. Put $\varphi : \mathcal{A} \to M_r(\mathbb{C})$ to be

$$\varphi(a) = \varphi_1 \oplus \varphi_2$$

and $\psi : M_r(\mathbb{C}) \to \mathcal{A}$ be

$$\psi(X) = t_1\psi_1(P_1XP_1) + t_2\psi_2(P_2XP_2),$$

where $P_1$ and $P_2$ are respectively the compressions from the natural embedding

$$M_{n(1)}(\mathbb{C}) \oplus M_{n(2)}(\mathbb{C}) \hookrightarrow M_r(\mathbb{C}).$$

Then $\psi \circ \varphi = t_1\theta_1 + t_2\theta_2$. $\square$

**13.2.11. Conditional Expectations.** Though we are not doing probability, this is a standard term in operator algebras:

**13.2.64. Definition.**

*Let $\mathcal{B} \subset \mathcal{A}$ be $C^*$-algebras and $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ a linear map. We say that*

- *$\mathcal{E}$ is a* ***projection*** *if $\mathcal{E}|_{\mathcal{B}} = \mathrm{id}_{\mathcal{B}}$;*
- *$\mathcal{E}$ is* ***$\mathcal{B}$-linear*** *if $\mathcal{E}(ab) = \mathcal{E}(a)b$ and $\mathcal{E}(ba) = b\mathcal{E}(a)$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$;*

- *$\mathcal{E}$ is a **conditional expectation** if it is a $\mathcal{B}$-linear projection which is also positive.*

The name *conditional expectation* should definitely raise eyebrows since—as we said—we are not doing any kind of probability. The connection is via the idea of conditional expectation of a subalgebra in probability. Namely, suppose $(X, \Sigma, \mu)$ is a **probability space** (a measure space where $\mu(X) = 1$) and $f : X \to \mathbb{R}^n$ is a bounded random variable, so $f \in L^\infty(X)$. If $\Sigma_0 \subset \Sigma$ is a subalgebra, the function $f$ may not be $\Sigma_0$-measurable. In this setting a conditional expectation of $f$ is a $\Sigma_0$-measurable function $f_0$ (so $f_0 \in L^\infty(\mu|_{\Sigma_0})$) such that $\int_E f_0\, d\mu = \int_E f\, d\mu$ for all $E \in \Sigma_0$. When we consider $L^\infty(\mu|_{\Sigma_0}) \subset L^\infty(\mu)$, the map $f \longmapsto f_0$ satisfies the conditions in Definition 13.2.64.

A rather direct observation is that if $\mathcal{A}$ is unital, then $\mathcal{B}$ is unital and $I_\mathcal{B} = \mathcal{E}(I_\mathcal{A})$. Indeed, for any $b \in \mathcal{B}$ we have

$$\mathcal{E}(I_\mathcal{A})b = \mathcal{E}(I_\mathcal{A} b) = \mathcal{E}(b) = b,$$

and similarly $b\mathcal{E}(I_\mathcal{A}) = b$ for all $b \in \mathcal{B}$.

*13.2.65. Example.* Let $\mathcal{A} = M_n(\mathbb{C})$, and $\mathcal{B}$ the diagonal subalgebra. Then the map $A \longmapsto \operatorname{diag}(A_{11}, \ldots, A_{nn})$ is a conditional expectation (Exercise 13.2.28).

*13.2.66. Example.* Let $\mathcal{A}$ be any unital C*-algebra and $\varphi \in \mathcal{A}^*$ a positive linear functional. Then $a \longmapsto \varphi(a)\, I_\mathcal{A}$ is a conditional expectation onto $\mathbb{C}\, I_\mathcal{A}$ (Exercise 13.2.29).

A more or less immediate result is

***13.2.67. Proposition.***

*A conditional expectation is completely positive.*

*Proof.* Let $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ be a conditional expectation. We may assume, without loss of generality, that $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$; this might require composing $\mathcal{E}$ with a $*$-isomorphism but that will not affect complete positivity. By shrinking $\mathcal{H}$ is needed, we may assume that $\overline{\mathcal{B}\mathcal{H}} = \mathcal{H}$. Now using Zorn's lemma we construct a maximal family of unit vectors $\{\xi_j\}$ such that $\mathcal{H} = \bigoplus_j \overline{\mathcal{B}\xi_h}$. Each of the subspaces in the direct sum is invariant for $\mathcal{E}(\mathcal{A}) = \mathcal{B}$. Then $\mathcal{E}$ is completely positive if and only if each of the maps $\mathcal{E}_j : \mathcal{A} \to \mathcal{B}(\overline{\mathcal{B}\xi_j})$ given by $\mathcal{E}_j(a) = \mathcal{E}(a)|_{\overline{\mathcal{B}\xi_j}}$ is completely positive (Exercise 13.2.14). So still without loss of generality we may assume that there exists $\xi \in \mathcal{H}$ with $\overline{\mathcal{B}\xi} = \mathcal{H}$.

Fix $a_1, \dots, a_n \in \mathcal{A}$, $b_1, \dots, b_n \in \mathcal{B}$.

$$\begin{aligned}\sum_{k,h=1}^{n} \langle \mathcal{E}(a_h^* a_k) b_k \xi, b_h \xi\rangle &= \sum_{k,h=1}^{n} \langle \mathcal{E}(b_h^* a_h^* a_k b_k)\xi, \xi\rangle \\ &= \Big\langle \mathcal{E}\Big(\sum_{k,h=1}^{n} b_h^* a_h^* a_k b_k\Big)\xi, \xi\Big\rangle \geq 0,\end{aligned}$$

since $\mathcal{E}$ is positive. The density of $\mathcal{B}\xi$ on $\mathcal{H}$ then gives us the same inequality for arbitrary vectors $\xi_1, \dots, \xi_n$. Then Lemma 13.2.10 yields that $\mathcal{E}$ is completely positive. □

Because a conditional expectation is a projection, we immediately get from the definition that $\|\mathcal{E}\| \geq 1$. But a lot more can be said.

***13.2.68. Proposition. (Tomiyama)***

*Let $\mathcal{B} \subset \mathcal{A}$ be $C^*$-algebras and $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ a linear map. The following statements are equivalent:*

*(i) $\mathcal{E}$ is a conditional expectation;*

*(ii) $\mathcal{E}$ is a linear projection that satisfies Kadison's Schwarz inequality (13.8);*

*(iii) $\mathcal{E}$ is a linear projection of norm 1.*

*Proof.* (i) $\implies$ (ii) Given $a \in \mathcal{A}$,

$$\begin{aligned}\mathcal{E}(a^*a) - \mathcal{E}(a)^*\mathcal{E}(a) &= \mathcal{E}(a^*a) - \mathcal{E}(a)^*\mathcal{E}(a) - \mathcal{E}(a)^*\mathcal{E}(a) + \mathcal{E}(a)^*\mathcal{E}(a) \\ &= \mathcal{E}(a^*a) - \mathcal{E}(\mathcal{E}(a)^*(a)) - \mathcal{E}(a^*\mathcal{E}(a)) + \mathcal{E}(a)^*\mathcal{E}(a) \\ &= \mathcal{E}(a^*a - \mathcal{E}(a)^*a - a^*\mathcal{E}(a) + \mathcal{E}(a)^*\mathcal{E}(a)) \\ &= \mathcal{E}\big((a - \mathcal{E}(a))^*(a - \mathcal{E}(a))\big) \geq 0.\end{aligned}$$

(ii) $\implies$ (iii) The inequality implies that $\mathcal{E}$ is positive. So from $a^*a \leq \|a\|^2 I_{\mathcal{A}}$ we have $\mathcal{E}(a^*a) \leq \|a\|^2 I_{\mathcal{B}}$. Then

$$\|\mathcal{E}(a)\|^2 = \|\mathcal{E}(a)^*\mathcal{E}(a)\| \leq \|\mathcal{E}(a^*a)\| \leq \|a\|^2 \|I_{\mathcal{B}}\| = \|a\|^2.$$

Thus $\|\mathcal{E}\| \leq 1$. Being a projection we also have that $\|\mathcal{E}\| \geq 1$, and thus $\|\mathcal{E}\| = 1$.

(iii) $\implies$ (i) Suppose first that $\mathcal{A} = \mathcal{M}$ and $\mathcal{B} = \mathcal{N}$ are von Neumann algebras. Fix $P \in \mathcal{M}$ a projection such that either $P \in \mathcal{N}$ or $Q = I_{\mathcal{M}} - P \in \mathcal{N}$. Then $PX, QX \in \mathcal{N}$ for all $X \in \mathcal{N}$. Given $X, Y \in \mathcal{M}$,

$$\|PX + QY\|^2 = \|(PX + QY)^*(PX + QY)\| = \|X^*PX + Y^*QY\|$$

$$\leq \|PX\|^2 + \|QY\|^2.$$

This should look unusual, having a "triangular inequality" for the square of the norm; and it is. For any $t \in \mathbb{R}$,

$$(1+t)^2\|Q\mathcal{E}(PX)\|^2 = \|Q\mathcal{E}(PX + tQ\mathcal{E}(PX)\|^2 \leq \|PX + tQ\mathcal{E}(PX)\|^2$$
$$\leq \|PX\|^2 + t^2\|Q\mathcal{E}(PX)\|^2.$$

We can rewrite this as $(1+2t)\|Q\mathcal{E}(PX)\|^2 \leq \|PX\|^2$ for all $t$, which forces $Q\mathcal{E}(PX) = 0$. This is $\mathcal{E}(PX) = P\mathcal{E}(PX)$. Exchanging the roles of $P$ and $Q$ we also get $P\mathcal{E}(QX) = 0$, which is $P\mathcal{E}(X) = P\mathcal{E}(PX)$. It follows that $P\mathcal{E}(X) = \mathcal{E}(PX)$ for all $X \in \mathcal{M}$. Taking adjoints we also have the equality on the right. We also have $I_{\mathcal{N}} = \mathcal{E}(I_{\mathcal{M}})$. Given any state $\varphi \in S(\mathcal{N})$ we have, using that $\|\mathcal{E}\| \leq 1$,

$$\|\varphi \circ \mathcal{E}\| \leq \|\varphi\|\,\|\mathcal{E}\| \leq 1.$$

As $\varphi \circ \mathcal{E}(I_{\mathcal{M}}) = \varphi(I_{\mathcal{N}}) = 1$, we get from Proposition 11.5.4 that $\varphi \circ \mathcal{E}$ is positive. As this can be done for all $\varphi \in S(\mathcal{N})$, from Corollary 11.5.10 we deduce that $\mathcal{E}$ is positive. As $\mathcal{E}$ is continuous (being bounded) and any von Neumann is spanned in norm by its projections (Corollary 12.4.16), the equalities $\mathcal{E}(XY) = \mathcal{E}(X)Y$ and $\mathcal{E}(YX) = Y\mathcal{E}(X)$ hold for any $X \in \mathcal{M}$ and $Y \in \mathcal{N}$.

When $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ and $\mathcal{A}, \mathcal{B}$ are C*-algebras, we consider the double adjoint $\mathcal{E}^{**} : \mathcal{A}^{**} \to \mathcal{B}^{**}$. It is a projection because $\mathcal{E}$ is, and it has norm 1 because $\mathcal{E}$ has, by Exercise 9.4.3. Because $\mathcal{B} \subset \mathcal{A}$ and there are natural weak*-continuous embeddings $\mathcal{A} \hookrightarrow \mathcal{A}^{**}$ and $\mathcal{B} \hookrightarrow \mathcal{B}^{**}$ we may assume that $\mathcal{B}^{**} \subset \mathcal{A}^{**}$ via Theorem 12.7.8. By the first part of the proof $\mathcal{E}^{**}$ is a conditional expectation onto $\mathcal{B}^{**}$, and then $\mathcal{E}$ is a conditional expectation as it is a restriction of one. ☐

*13.2.69. Remark*. By (iii) in Proposition 9.9.6 and Arveson's Extension Theorem, if $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ is a C*-algebra and there exists a conditional expectation $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{A}$, then $\mathcal{A}$ is injective. And by (ii) in Proposition 9.9.6, together with Tomiyama (Proposition 13.2.68) the converse also holds: if $\mathcal{A}$ is injective, then there exists a conditional expectation $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{A}$.

*13.2.70. Example*. Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ be a maximal abelian von Neumann algebra which is also **atomic**; this means that for every projection $P \in \mathcal{A}$ there exists a projection $P_0 \leq P$ with $P_0 \in \mathcal{A}$ minimal (that, is $P_0\mathcal{A} = \mathbb{C}\,P_0$). An example of this is $\ell^\infty(\mathbb{N}) \subset \mathcal{B}(\ell^2(\mathbb{N}))$; this is a masa by Proposition 12.5.10 since it is abelian and the vector $\sum_n \frac{1}{n}\,e_n$ is cyclic for it (we can also use Proposition 12.5.8). Via Zorn we construct a maximal family $\{F_j\}$ of pairwise orthogonal minimal projections in $\mathcal{A}$. The maximality guarantees that $\sum_j F_j = I_{\mathcal{A}} = I_{\mathcal{H}}$. Because $\mathcal{A}$

is maximal abelian, each minimal projection in $\mathcal{A}$ is actually minimal in $\mathcal{B}(\mathcal{H})$, for otherwise we could halve one such projection and create a larger abelian algebra than contains $\mathcal{A}$. Let $\mathcal{E}:\mathcal{B}(\mathcal{H})\to\mathcal{A}$ be given by

$$\mathcal{E}(T)=\sum_j F_jTF_j.$$

Then $\mathcal{E}$ is a faithful normal conditional expectation (Exercise 13.2.31).

*13.2.71. Example*. There are other kind of maximal abelian algebras, which have no minimal projections. For example consider $\mathcal{A}=L^\infty[0,1]\subset\mathcal{B}(L^2[0,1])$; this is a masa by Proposition 12.5.10 since it is abelian and the constant function 1 is a cyclic vector for it (as above, we can also use Proposition 12.5.8). To exploit the group nature of the business, we may change the inclusion (via a unitary) to $L^\infty(\mathbb{T})\subset L^2(\mathbb{T})$. Since the unitaries $\{e^{int}\}_{n\in\mathbb{Z}}$ generate $C(\mathbb{T})$ as a Banach space (by Theorem 8.1.11), they generate $L^\infty(\mathbb{T})$ as a von Neumann algebra via an application of dominated convergence.

On $\ell^\infty(\mathbb{Z})$ we consider the states

$$\varphi_N(a)=\frac{1}{2N+1}\sum_{n=-N}^{N}a(n).$$

Since the state space of $\ell^\infty(\mathbb{Z})$ is weak$^*$-compact (Proposition 11.5.15), there exists an accumulation point $\varphi$ of $\{\varphi_N\}_N$. This is what is known as an **invariant mean**, in the sense that if $b(n)=a(n+m)$ for all $n$ (that is, $b$ is a translate of $a$), then $\varphi(b)=\varphi(a)$; this follows from writing $\varphi$ as the limit of a net $\{\varphi_{N_j}\}$ (Exercise 13.2.32).

For $T\in\mathcal{B}(\mathcal{H})$ define by $f_T(n)=M_{e^{int}}TM_{e^{int}}^*$, $n\in\mathbb{Z}$. Consider the form $[\xi,\eta]=\varphi(f_{T,\xi,\eta})$, where $f_{T,\xi,\eta}\in\ell^\infty(\mathbb{Z})$ is the function $f_{T,\xi,\eta}(n)=\langle f_T(n)\xi,\eta\rangle$. The form is sequilinear by construction and, since $\|f_T(n)\|=\|T\|$, we get that $[\xi,\eta]\leq\|T\|\,\|\xi\|\,\|\eta\|$. By Proposition 10.1.5 there exists a unique $\mathcal{E}(T)\in\mathcal{B}(\mathcal{H})$ such that $\|\mathcal{E}(T)\|\leq\|T\|$ (since $\|T\|$ is an upper bound for the norm of the form) and

$$\langle\mathcal{E}(T)\xi,\eta\rangle=\varphi(f_{T,\xi,\eta}),\qquad \xi,\eta\in\mathcal{H}.$$

The uniqueness forces $\mathcal{E}$ to be linear. If $A\in\mathcal{A}$, then $f_T(A)=A$ and then $\mathcal{E}(A)=A$. Conversely, for any $T\in\mathcal{B}(\mathcal{H})$ and $m\in\mathbb{Z}$ consider the function $n\longmapsto M_{e^{imt}}f_T(n)M_{e^{imt}}^*=f_T(n+m)$. It follows that the corresponding function $g_{\xi,\eta}(n)=\langle f_T(n+m)\xi,\eta\rangle$ is a translate of $f_T$. Then

$$\begin{aligned}\langle M_{e^{imt}}\mathcal{E}(T)M_{e^{imt}}^*\xi,\eta\rangle&=\varphi(f_{T,M_{e^{imt}}\xi,M_{e^{imt}}\eta})=\varphi(g_{\xi,\eta})=\varphi(f_{T,\xi,\eta})\\&=\langle\mathcal{E}(T)\xi,\eta\rangle.\end{aligned}$$

As this can be done for all $\xi, \eta \in \mathcal{H}$ we get that $\mathcal{E}(T) \in \mathcal{A}' = \mathcal{A}$. We have shown that $\mathcal{E}$ is a projection of norm 1 onto $\mathcal{A}$, and thus a conditional expectation by Proposition 13.2.68.

The argument above works in way more generality, for any von Neumann algebra of the form $\mathcal{M} = G'$ with $G$ **amenable**; though we will not discuss it here, amenability is a big deal for groups and constructions with them in the world of operator algebras. While we will not deal with it here, it is worth being familiar with; we refer to [Dav25, Section 9.13].

The reason we are working on this example after having Example 13.2.70 is that as opposed to that example, the expectation here can never be normal. Indeed, fix $K \in \mathcal{K}(L^2[0,1])$ and for $t \in [0,1]$ and $\delta > 0$ consider the projections $P_\delta = M_{1_{[t-\delta,t+\delta]}}$. We have $\|P_\delta K\| \to 0$ by Exercise 12.4.9. Given $\varepsilon > 0$ for each $t$ there exists $\delta(t)$ with $\|P_{\delta(t)} K\| < \varepsilon$. By the compactness of $[0,1]$ there exists a partition $\{t_k\}_{k=0}^n$ of $[0,1]$ with $\|P_{[t_{k-1},t_k]} K\| < \varepsilon$. As these projections are in $\mathcal{A}$,

$$\|P_{[t_{k-1},t_k]} \mathcal{E}(K)\| = \|\mathcal{E}(P_{[t_{k-1},t_k]} K)\| \leq \|P_{[t_{k-1},t_k]} K\| < \varepsilon.$$

But $\mathcal{E}(K) = \sum_k P_{[t_{k-1},t_k]} \mathcal{E}(K)$, which gives $\|\mathcal{E}(K)\| < \varepsilon$ as the projections are pairwise orthogonal. As $\varepsilon$ was arbitrary, $\mathcal{E}(K) = 0$. This means that $\mathcal{E}(K) = 0$ for all compact $K$ while $\mathcal{E}(I_\mathcal{H}) = I_\mathcal{H}$, and therefore $\mathcal{E}$ is not normal.

We will see other examples of conditional expectations further down the text. In particular there are natural conditional expectations arising from tensor and crossed products onto certain canonical subalgebras. Many von Neumann algebras have conditional expectations onto some of their subalgebras, and conditional expectations are staples of von Neumann algebra works. Landmarks to mention are Takesaki's characterization of the existence of a conditional expectation onto a subalgebra [Tak72] and Connes' monumental work on the classification of injective factors [Con76]. Later conditional expectations played a crucial role in Jones' Index Theory (started on [Jon83], see also [PP86] for an example). Modern von Neumann algebra research is majorly about subfactors (grown out of Jones' Index Theory) and Popa's deformation/rigidity theory (grown out of [Pop06a, Pop06b, Pop07]).

**13.2.12. Exercises.**

(13.2.1) Prove that if $\mathcal{S} \subset \mathcal{A}$ is an operator system, then $M_n(\mathcal{S}) \subset M_n(\mathcal{A})$ is an operator system.

(13.2.2) Show that that if $n, m \in \mathbb{N}$ with $n < m$, and $\phi : \mathcal{S} \to \mathcal{B}$, then $\|\phi^{(m)}\| \geq \|\phi^{(n)}\|$, and $\phi^{(m)} \geq 0$ implies $\phi^{(n)} \geq 0$.

(13.2.3) Let $\mathcal{S} = \mathcal{A}$ be a C$^*$-algebra, and $\phi : \mathcal{A} \to \mathcal{B}$ a $*$-homomorphism. Show that $\phi$ is completely positive.

(13.2.4) Show an example of a C$^*$-algebra $\mathcal{A}$ with a dense subalgebra $\mathcal{A}_0$ and a $*$-homomorphism $\rho : \mathcal{A}_0 \to \mathcal{A}_0$ that is unbounded (so, in particular, it doesn't extend to $\mathcal{A}$).

(13.2.5) Given a compact Hausdorff space $T$, show that the C*-algebras $\mathcal{A} = M_n(C(T))$ and $\mathcal{B} = C(T, M_n(\mathbb{C}))$ are canonically isomorphic, where the norm in $\mathcal{B}$ is given by

$$\|y\|_{\mathcal{B}} = \sup\{\|y(t)\| : \ t \in T\}.$$

(13.2.6) Prove that a positive map $\phi : \mathcal{S} \to \mathcal{B}$ maps selfadjoint elements to selfadjoint elements.

(13.2.7) Show that

$$\alpha + \beta z + \gamma \bar{z} \geq 0, \ z \in \overline{\mathbb{D}} \iff \alpha \geq 0, \ \gamma = \bar{\beta}, \ 2|\beta| \leq \alpha$$
$$\iff \begin{bmatrix} \alpha & 2\beta \\ 2\gamma & \alpha \end{bmatrix} \geq 0.$$

(13.2.8) Prove equation (13.4), i.e.

$$\sum_{k,j=1}^{n} \langle \phi(a_j^* a_k)\xi_k, \xi_j \rangle = \left\langle \phi^{(n)}(A^*A)\xi, \xi \right\rangle,$$

where $\xi = (\xi_1, \dots, \xi_n)^\top \in \mathcal{H}^n$ and $A \in M_n(\mathcal{A})$ is the matrix with some row $a_1, \dots, a_n$ and zeroes elsewhere.

(13.2.9) Show that if $X \in M_n(\mathcal{A})$ then we can write $X = X_1 + \cdots + X_n$—where $X_k$ is the matrix such that its $k^{\text{th}}$ row is that of $X$, and the rest of the rows are zero—and then $X^*X = \sum_k X_k^* X_k$.

(13.2.10) In the proof of Proposition 13.2.12, show that $\|X\| = \|\xi\|$.

(13.2.11) Show that if $A \in M_n(\mathbb{C})$ and $B \in M_n(\mathcal{S})$, then $AB \in M_n(\mathcal{S})$. Show an example where $A, B \in M_n(\mathcal{S})$ and $AB \notin M_n(\mathcal{S})$.

(13.2.12) Show that the set $K_{00}$ in the proof of Stinespring's Theorem 13.2.15 is actually a subspace. *(Hint: Cauchy-Schwarz)*

(13.2.13) Prove the non-unital case and the uniqueness, up to unitary conjugation, of the Stinespring's Dilation (Theorem 13.2.15).

(13.2.14) Let $\Phi : \mathcal{A} \to \mathcal{B}$ be a linear map. Let $\gamma : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ be a faithful representation and $\mathcal{H} = \bigoplus_j \mathcal{H}_j$ a decomposition of $\mathcal{H}$ such that $\gamma(\mathcal{B})\mathcal{H}_j \subset \mathcal{H}_j$. Show that $\Phi$ is completely positive if and only if each restriction $P_j(\gamma \circ \Phi)P_j$ is completely positive, where $P_j : \mathcal{H} \to \mathcal{H}_j$ is the canonical projection.

(13.2.15) Show that the transpose map $\phi_T : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ is positive, unital, and contractive.

(13.2.16) Work out the missing details in Example 13.2.20. Namely, show that the matrix inside $\phi_T^{(2)}$ is positive, while its image is not positive.

(13.2.17) Let $\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be contractive and completely positive with minimal Stinespring triple $(\pi, \mathcal{K}, V)$. Use ideas from the proof of Lemma 13.2.37 to show there exists a $*$-homomorphism $\rho : \varphi(\mathcal{A})' \to \pi(\mathcal{A})' \subset \mathcal{B}(\mathcal{K})$ that satisfies

$$\varphi(a)T = V^*\pi(a)\rho(T)V, \qquad a \in \mathcal{A}, \ T \in \varphi(A)'.$$

(13.2.18) Let $X \in M_n(\mathcal{A})^+$. Show that if $X_{kk} = 0$ for some $k$, then $X_{kj} = X_{jk} = 0$ for all $j = 1, \ldots, n$.

(13.2.19) Let $\begin{bmatrix} x & y \\ y^* & z \end{bmatrix} \in M_2(\mathcal{A})$ with $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} x & y \\ y^* & z \end{bmatrix} \leq \begin{bmatrix} I_{\mathcal{A}} & 0 \\ 0 & 0 \end{bmatrix}$. Show that $y = z = 0$. Use the same idea to conclude that if $X \in M_n(\mathcal{A})$ and $0 \leq X \leq I_{\mathcal{A}} \otimes E_{kk}$, then $X = a \otimes E_{kk}$ for some $a \in \mathcal{A}$.

(13.2.20) Let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be unital and 2-positive. Show that $\phi$ is bounded and $\|\phi\| = 1$.

(13.2.21) Let $\mathcal{A}$ be a non-unital C$^*$-algebra and $f \in S(\mathcal{A})$ pure. Show that the unique extension $\tilde{f}$ of $f$ to $\tilde{\mathcal{A}}$ (which exists by Proposition 11.5.6) is pure.

(13.2.22) Show that, using the matrix units $\{E_{kj}\}$ as the basis of $M_n(\mathbb{C})$, the basis can be ordered in such a way that the matrix representation of the multiplication operator $M_X : Y \longmapsto XY$ is $X \otimes I_n$.

(13.2.23) Let $\{\phi_j\} \subset CP(\mathcal{S}, \mathcal{B}(\mathcal{H}))$ be a bounded net. Show that $\phi_j \xrightarrow{BW} \phi$ if and only if $\langle \phi_j(x)\xi, \xi\rangle \to \langle \phi(x)\xi, \xi\rangle$ for all $\xi \in \mathcal{H}$ and $x \in \mathcal{S}$.

(13.2.24) Let $P \in \mathcal{B}(\mathcal{H})$ be a finite-rank projection, $n = \operatorname{Tr}(P)$. Show that $P\mathcal{B}(\mathcal{H})P \simeq M_n(\mathbb{C})$ as C$^*$-algebras.

(13.2.25) Show that the composition of cp maps is cp.

(13.2.26) Show that if $A \in M_n(\mathcal{S})$, i.e. $A = \sum_{k,j=1}^n a_{kj} \otimes E_{kj}$ with $a_{kj} \in \mathcal{S}$ for all $k, j$, then $\|A\| \leq \sum_{k,j} \|a_{kj}\|$.

(13.2.27) Let $\mathcal{A}$ be a C$^*$-algebra and $\varphi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a contractive completely positive map. Show that $\varphi$ admits a unique ucp extension to the unitization $\tilde{\mathcal{A}}$.

(13.2.28) Let $\mathcal{A} = M_n(\mathbb{C})$, and $\mathcal{B}$ the diagonal subalgebra. Show that the map $A \longmapsto \operatorname{diag}(A_{11}, \ldots, A_{nn})$ is a conditional expectation.

(13.2.29) Let $\mathcal{A}$ be a unital C$^*$-algebra and $\varphi \in \mathcal{A}^*$ a positive linear functional. Show that $a \longmapsto \varphi(a)\, I_{\mathcal{A}}$ is a conditional expectation onto $\mathbb{C}\, I_{\mathcal{A}}$.

(13.2.30) Write a direct proof of (ii) $\Longrightarrow$ (i) in Proposition 13.2.68.

(13.2.31) Show that the map $\mathcal{E}$ from Example 13.2.70 is a faithful normal conditional expectation.

(13.2.32) Show that if on $\ell^\infty(\mathbb{Z})$ we consider the states

$$\varphi_N(a) = \frac{1}{2N+1} \sum_{n=-N}^{N} a(n).$$

and $\varphi$ is a weak$^*$-accumulation point for $\{\varphi_N\}_N$, then $\varphi$ is an invariant mean. That is, if $b(n) = a(n+m)$ for all $n$ (that is, $b$ is a translate of $a$), show that $\varphi(b) = \varphi(a)$.

## 13.3. Group Algebras

During the whole process of building theory for C*-algebras, examples were not overabundant. This should not be interpreted as evidence that C*-algebras are not too varied. After all $C[0,1]$, $\ell^\infty(\mathbb{N})$, $M_n(\mathbb{C})$, and $\mathcal{B}(\ell^2(\mathbb{N}))$ feel so radically different that this should be taken as an indication that lots of C*-algebras are out there. Also, $\mathcal{B}(\mathcal{H})$ is non-separable as a C*-algebra when $\dim \mathcal{H} = \infty$, so infinite-dimensional C*-algebras which are not $\mathcal{B}(\mathcal{H})$ can be constructed by getting the C*-algebra generated by finite/countable subsets of $\mathcal{B}(\mathcal{H})$. Here we will examine another way to get C*-algebras, which is through groups.

Given a group $G$, we get a natural algebra $\mathbb{C}G$ by considering formal linear combinations of elements of $G$ and using the group operation to induce the product. We can make it a $*$-algebra by prescribing $g^* = g^{-1}$. So if we were able to define a C*-norm on $\mathbb{C}G$, the completion would be a C*-algebra. This can be done in several ways, as we will see.

**13.3.1. The Left and Right Regular Representations.** One natural way to put a norm on $\mathbb{C}G$ is to embed it in some $\mathcal{B}(\mathcal{H})$. At first glance it is not clear how to see elements of an abstract group as operators on a Hilbert space. But we have already been faced with a similar situation, when we had an abstract C*-algebra and we wanted to see it as operators on a Hilbert space. In the GNS construction we managed to build a Hilbert space out of the algebra itself, and use the algebra multiplication as the action on the space. Here we will do something similar. Historically, the construction we are about to undertake is older than the GNS construction, and it probably inspired it.

**13.3.1. Definition.**

*Let $G$ be a group with unit $e$. Put $\mathcal{H}_G = \ell^2(G)$. We write $\delta_g$ for the element of the canonical basis corresponding to $g \in G$. The **left regular representation** is the map*

$$\lambda : G \to \mathcal{B}(\mathcal{H}_G) \qquad \text{given by} \qquad (\lambda(g)\xi)(h) = \xi(g^{-1}h),\ \xi \in \mathcal{H}_G,\ h \in G.$$

Of course one needs to check that $\lambda(g)$ is indeed a bounded operator. We summarize below the most basic properties of the left regular representation.

**13.3.2. Proposition.**

*Let $G$ be a group and $\lambda : G \to \mathcal{H}_G$ its left regular representation. Let $g, h \in G$. Then*

(i) $\lambda(g) \in \mathcal{B}(\mathcal{H}_G)$;

(ii) $\lambda(g)\delta_h = \delta_{gh}$;

(iii) $\lambda(g)$ *is a unitary and* $\lambda(g)^* = \lambda(g^{-1})$;

(iv) $\lambda : G \to \mathcal{U}(\mathcal{H}_G)$ *is a group homomorphism.*

*Proof.* Exercise 13.3.1. □

The property $\lambda(g)\delta_h = \delta_{gh}$ shows that we may somehow think of $\lambda(g)$ as a bilateral shift.

Since groups are not abelian in general, everything we did we might repeat on the right:

**13.3.3. Definition.**

*The* ***right regular representation*** *is the map*

$$\rho : G \to \mathcal{B}(\mathcal{H}_G) \qquad \text{given by} \qquad (\rho(g)\gamma)(h) = \gamma(hg), \ \gamma \in \mathcal{H}_G, \ h \in G.$$

As before,

**13.3.4. Proposition.**

*Let $G$ be a group and $\rho : G \to \mathcal{H}_G$ its right regular representation. Let $g, h \in G$. Then*

(i) $\rho(g) \in \mathcal{B}(\mathcal{H}_G)$;

(ii) $\rho(g)\delta_h = \delta_{hg^{-1}}$;

(iii) $\rho(g)$ *is a unitary and* $\rho(g)^* = \rho(g^{-1})$;

(iv) $\rho : G \to \mathcal{B}(\mathcal{H}_G)$ *is a group homomorphism.*

*Proof.* Exercise 13.3.2. □

The apparent asymmetry from $g$ to $g^{-1}$ in the definition of the left and right regular representations stems from the fact that we want them to be group homomorphisms.

**13.3.2. The Reduced C*-algebras.** Now that we have a way to see our group $G$ as operators in some $\mathcal{B}(\mathcal{H})$, we can look at the C*-algebra they generate. We will talk about *discrete groups* to emphasize that we consider no topology on them. Besides the reduced C*-algebras mentioned below, one can also consider the **full C*-algebra** of the group; we defer this to Exercise 13.3.10.

**13.3.5. Definition.**

*Let $G$ be a discrete group. The **(left) reduced C*-algebra** of $G$ is*

$$C^*_\lambda(G) = C^*(\lambda(G)) \subset \mathcal{B}(\mathcal{H}_G).$$

*Similarly, the **right reduced C*-algebra** of $G$ is*

$$C^*_\rho(G) = C^*(\rho(G)) \subset \mathcal{B}(\mathcal{H}_G).$$

*A linear functional $\varphi$ on a C*-algebra is **tracial** if $\varphi(xy) = \varphi(yx)$ for all $x, y$.*

*The **canonical trace** $\tau$ on the reduced C*-algebras is the linear functional induced by*

$$\tau(\lambda(g)) = \tau(\rho(g)) = \begin{cases} 1, & g = e \\ 0, & g \neq e \end{cases}$$

*The **canonical unitary** $J$ is the map given by $(J\xi)(g) = \xi(g^{-1})$.*

When mentioned without the left/right qualifier, the standard is to assume that one is talking about the left reduced C*-algebra. The notation $C^*_r(G)$ is common in the literature, where the $r$ refers to "reduced" and not to "right".

The map $J$ is always a unitary and $J\lambda(g)J = \rho(g)$ for all $g \in G$ (Exercise 13.3.4).

A basic observation that we will use a few times is that $xy = yx$ for all $x \in C^*_\lambda(G)$ and $y \in C^*_\rho(G)$ (Exercise 13.3.3).

When looking at $C^*_\lambda(G)$ as a single algebra, there are two equivalent points of view here. One can think of $C^*_\lambda(G)$ as an algebra of operators on the Hilbert space $\mathcal{H}_G$. The other point of view is that you can extend $\lambda$ by linearity to a representation of $\mathbb{C}G$, and then define a norm on $\mathbb{C}G$ by $\|x\|_\lambda = \|\lambda(x)\|_{\mathcal{B}(\mathcal{H}_G)}$. One then considers $C^*_\lambda(G)$ as the abstract completion of $\mathbb{C}G$ with respect to the norm $\|\cdot\|_\lambda$.

***13.3.6. Proposition.***

*Let $G$ be a group. The canonical trace is a bounded faithful linear functional on $C^*_\lambda(G)$, that satisfies*

*(i) $\tau(xy) = \tau(yx)$ for all $x, y \in C^*_\lambda(G)$;*

*(ii)* $\tau(x) = \langle x\delta_e, \delta_e \rangle$.

*The same properties hold on $C^*_\rho(G)$.*

*Proof.* We have

$$\langle \lambda(g)\delta_e, \delta_e \rangle = \langle \delta_g, \delta_e \rangle = \begin{cases} 1, & g = e \\ 0, & g \neq e \end{cases}$$

So $\tau(x) = \langle x\delta_e, \delta_e \rangle$ for all $x \in \mathbb{C}\lambda(G)$. As the linear functional $T \longmapsto \langle T\delta_e, \delta_e \rangle$ is bounded (since $|\langle T\delta_e, \delta_e \rangle| \leq \|T\|$), $\tau$ extends uniquely to all of $C^*_\lambda(G)$ as a bounded linear functional.

Let $x = \sum_{j=1}^m \alpha_j \lambda(g_j)$ and $y = \sum_{k=1}^n \beta_k \lambda(h_k)$ (with the standard convention that the $g_j$ are pairwise distinct and that the $h_k$ are pairwise distinct). Then

$$\begin{aligned} \tau(xy) &= \sum_{k,j} \alpha_j \beta_k \langle \lambda(g_j h_k)\delta_e, \delta_e \rangle = \sum_{k,j} \alpha_j \beta_k \langle \delta_{g_j h_k}, \delta_e \rangle \\ &= \sum_{h_k = g_j^{-1}} \alpha_j \beta_k = \sum_{k,j} \alpha_j \beta_k \langle \delta_{h_k g_j}, \delta_e \rangle = \tau(yx). \end{aligned}$$

The equality then extends to all of $C^*_\lambda(G)$ by boundedness. For faithfulness, if $\tau(x^*x) = 0$ then $\|x\delta_e\| = 0$, so $x\delta_e = 0$ and hence $x\delta_g = \rho(g^{-1})x\delta_e = 0$ for all $g \in G$; thus $x = 0$.

For $C^*_\rho(G)$ the computations are entirely analogous. $\square$

*13.3.7. Remark*. Since $C^*_\lambda(G)$ and $C^*_\rho(G)$ commute and $\tau$ is tracial on both, it is tempting to think that $\tau$ is tracial on the C*-algebra generated by $C^*_\lambda(G)$ and $C^*_\rho(G)$. But that is not the case unless $G$ is abelian. Indeed, if $G$ is not abelian let $a, b \in G$ such that $ab \neq ba$. Then if $S = \lambda(ab^{-1})$ and $T = \lambda(b)\rho(a)$, we have that $S, T \in \operatorname{alg}\left\{C^*_\lambda(G) \cup C^*_\rho(G)\right\}$ and

$$\tau(ST) = \tau(\lambda(ab^{-1})\lambda(b)\rho(a)) = \tau(\lambda(a)\rho(a)) = \langle \lambda(a)\delta_e, \rho(a^{-1})\delta_e \rangle = \langle \delta_a, \delta_a \rangle = 1,$$

while

$$\begin{aligned} \tau(TS) &= \tau(\lambda(b)\rho(a)\lambda(ab^{-1})) = \tau(\lambda(bab^{-1})\rho(a)) \\ &= \langle \lambda(bab^{-1})\delta_e, \rho(a^{-1})\delta_e \rangle = \langle \delta_{bab^{-1}}, \delta_a \rangle = 0. \end{aligned}$$

We will need the following two lemmas in Section 13.4.

**13.3.8. Lemma.**

*Let $G$ be a discrete group, $F \subset E \subset G$ subsets such that there exist $a_1, a_2 \in G$ with $F \cup a_1 F a_1^{-1} = E$, and $F$, $a_2^{-1} F a_2$, and $a_2 F a_2^{-1}$ disjoint subsets of $E$. Let $\varepsilon > 0$ and $\xi \in \mathcal{H}_G$ such that, for $j = 1, 2$,*

$$\sum_g |\xi(a_j g a_j^{-1}) - \xi(g)|^2 < \varepsilon^2. \tag{13.14}$$

*Then $\|P_E \xi\| \leq 14\varepsilon$, where $P_E$ is the orthogonal projection onto the subspace of elements with support in $E$.*

*Proof.* By (13.14) and the reverse triangle inequality,

$$\begin{aligned} \big|\, \|P_{a_1 F a_1^{-1}} \xi\| - \|P_F \xi\| \,\big| &= \left| \left( \sum_{g \in F} |\xi(a_1 g a_1^{-1})|^2 \right)^{1/2} - \left( \sum_{g \in F} |\xi(g)|^2 \right)^{1/2} \right| \\ &\leq \left( \sum_{g \in F} |\xi(a_1 g a_1^{-1}) - \xi(g)|^2 \right)^{1/2} < \varepsilon. \end{aligned}$$

Let $s = \|P_E \xi\|$. Then

$$\begin{aligned} \big|\, \|P_{a_1 F a_1^{-1}} \xi\|^2 - \|P_F \xi\|^2 \,\big| &= \Big( \big|\, \|P_{a_1 F a_1^{-1}} \xi\| - \|P_F \xi\| \,\big| \Big) \Big( \|P_{a_1 F a_1^{-1}} \xi\| + \|P_F \xi\| \Big) \\ &\leq 2s\varepsilon. \end{aligned}$$

Hence $\|P_{a_1 F a_1^{-1}} \xi\|^2 \leq \|P_F \xi\|^2 + 2s\varepsilon$. From this (note that $P_E = P_{a_1 F a_1^{-1}} + P_F$),

$$s^2 = \|P_E \xi\|^2 = \|P_{a_1 F a_1^{-1}} \xi\|^2 + \|P_F \xi\|^2 \leq 2(\|P_F \xi\|^2 + s\varepsilon),$$

which implies

$$\|P_F \xi\|^2 \geq \frac{s^2}{2} - s\varepsilon.$$

With $a_2$ and $a_2^{-1}$ in place of $a_1$ we get

$$\big|\, \|P_{a_2 F a_2^{-1}} \xi\|^2 - \|P_F \xi\|^2 \,\big| \leq 2s\varepsilon, \qquad \big|\, \|P_{a_2^{-1} F a_2} \xi\|^2 - \|P_F \xi\|^2 \,\big| \leq 2s\varepsilon.$$

We can then write

$$\|P_{a_2 F a_2^{-1}} \xi\|^2 \geq \|P_F \xi\|^2 - 2s\varepsilon, \qquad \|P_{a_2^{-1} F a_2} \xi\|^2 \geq \|P_F \xi\|^2 - 2s\varepsilon.$$

Therefore

$$\begin{aligned} s^2 = \|P_E \xi\|^2 &\geq \|P_F \xi\|^2 + \|P_{a_2 F a_2^{-1}} \xi\|^2 + \|P_{a_2^{-1} F a_2} \xi\|^2 \\ &\geq 3\|P_F \xi\|^2 - 4s\varepsilon \geq \frac{3s^2}{2} - 3s\varepsilon - 4s\varepsilon = \frac{3s^2}{2} - 7s\varepsilon. \end{aligned}$$

Thus

$$\|P_E \xi\| = s \leq 14\varepsilon. \qquad \square$$

The next lemma is specific for the free group $\mathbb{F}_2$ on two generators.

***13.3.9. Lemma.***

*Let $a_1, a_2$ be the generators of $\mathbb{F}_2$ and let $T \in \mathcal{B}(\mathcal{H}_{\mathbb{F}_2})$ be given by*

$$T = \frac{1}{4}\big(\lambda(a_1)\rho(a_1) + \lambda(a_2)\rho(a_2) + \lambda(a_1^{-1})\rho(a_1^{-1}) + \lambda(a_2^{-1})\rho(a_2^{-1})\big).$$

*Then $T$ is selfadjoint, $\|T\| = 1$, $T\delta_e = \delta_e$, $\sigma(T|_{(\mathbb{C}\delta_e)^\perp}) \subset [-1, 1-\delta]$, where $\delta = 40^{-4}$, and $\sigma(T) \subset [-1, 1-\delta] \cup \{1\}$ .*

*Proof.* Recalling that $\lambda$ and $\rho$ commute, each term $\lambda(a_j)\rho(a_j) + \lambda(a_j^{-1})\rho(a_j^{-1})$ is selfadjoint, and hence $T$ is selfadjoint. The equality $T\delta_e = \delta_e$ is straightforward. In particular $\mathbb{C}\delta_e$ is invariant for $T$, and so $(\mathbb{C}\delta_e)^\perp$ is also invariant for $T$ (since $T$ is selfadjoint). Let $\xi \in (\mathbb{C}\delta_e)^\perp$ with $\|\xi\| = 1$. Put $\varepsilon = \|\xi - T\xi\|$. We can see $T$ as a double average. Then Exercise 4.2.9, applied twice, gives us

$$\begin{aligned}\|\xi - \lambda(a_1)\rho(a_1)\xi\| &\le 2\Big\|\xi - \frac{1}{2}\big(\lambda(a_1)\rho(a_1)\xi + \lambda(a_2)\rho(a_2)\xi\big)\Big\|^{1/2} \\ &\le 2\,2^{1/2}\,\|\xi - T\xi\|^{1/4} = 2^{3/2}\,\varepsilon^{1/4}.\end{aligned}$$

And, similarly,

$$\|\xi - \lambda(a_2)\rho(a_2)\xi\| \le 2^{3/2}\,\varepsilon^{1/4}.$$

We have

$$\sum_g |\xi(g) - \xi(a_1 g a_1^{-1})|^2 = \|\xi - \lambda(a_1)\rho(a_1)\xi\|^2 \le (2^{3/2}\varepsilon^{1/4})^2.$$

If we take $E = \mathbb{F}_2 \setminus \{e\}$ and $F = \{w \in \mathbb{F}_2 : \ w = a_1^{-1}\cdots\} \setminus \{a_1^{-1}\}$, then Lemma 13.3.8 gives us

$$1 = \|\xi\| = \|P_E\xi\| \le 14 \cdot 2^{3/2}\,\varepsilon^{1/4} \le 40\varepsilon^{1/4}.$$

Hence $\varepsilon \ge 40^{-4}$. Let $\delta = 40^{-4}$. When $1 - \delta < \lambda < 1$,

$$\|(T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}})\xi\| + |1 - \lambda|\,\|\xi\| = \|T\xi - \lambda\xi\| + \|\lambda\xi - \xi\| \ge \|T\xi - \xi\| = \varepsilon.$$

As $1 - \lambda < \delta \le \varepsilon$,

$$\|(T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}})\xi\| \ge (\varepsilon - (1 - \lambda))\|\xi\|.$$

We will see that $T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}}$ is bounded below. Indeed, if $\eta \in \mathcal{H}_{\mathbb{F}_2}$ we can write $\eta = \alpha\delta_e + \beta\xi$, with $\xi \in (\mathbb{C}\delta_e)^\perp$ a unit vector. Thus, as both $\mathbb{C}\delta_e$ and $(\mathbb{C}\delta_e)^\perp$ are invariant for $T$,

$$\begin{aligned}\|(T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}})\eta\|^2 &= |\alpha|^2\,\|(T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}})\delta_e\|^2 + |\beta|^2\|(T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}})\xi\|^2 \\ &\ge |\alpha|^2(1 - \lambda)^2 + \beta|^2(\varepsilon - (1 - \lambda))| \ge c(|\alpha|^2 + |\beta|^2) = c\|\eta\|^2,\end{aligned}$$

where $c = \min\{1-\lambda, \varepsilon - (1-\lambda)\}$. Therefore $T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}}$ is selfadjoint and bounded below, so its range is dense and closed, and thus it is bijective. It follows that $T - \lambda I_{\mathcal{H}_{\mathbb{F}_2}}$ is invertible by the Inverse Mapping Theorem. This means that $\lambda \notin \sigma(T)$. As $\|T\| \leq 1$ by construction and $T\delta_e = \delta_e$, we have $\sigma(T) \subset [-1,1] \setminus (1-\delta, 1)$. That is, $\sigma(T) \subset [-1, 1-\delta] \cup \{1\}$. ☐

The next result, due to Powers [Pow75], has been greatly studied and generalized. We refer to [Ken20] as a starting point for older and newer bibliography in the subject. The proof we offer follows [PS79].

***13.3.10. Theorem. (Powers)***

*The $C^*$-algebra $C^*_\lambda(\mathbb{F}_2)$ is simple.*

*Proof.* We consider $\mathcal{H} = \mathcal{H}_{\mathbb{F}_2}$ and $\mathcal{A} = C^*_\lambda(\mathbb{F}_2) \subset \mathcal{B}(\mathcal{H})$. We know from Proposition 13.3.6 that $\mathcal{A}$ has a faithful tracial functional $\tau$, given by $\tau(x) = \langle \lambda(x)\delta_e, \delta_e \rangle$. The result will be a consequence of the following

**Claim.** Let $S \in \mathcal{A}^{\mathrm{sa}}$. Then

$$\tau(S)\, I_{\mathcal{H}} \in \overline{\mathrm{conv}}\{\lambda(g)S\lambda(g)^* : \ g \in \mathbb{F}_2\}. \tag{13.15}$$

Assuming the Claim, suppose $\mathcal{J} \subset \mathcal{A}$ is an ideal and $S \in \mathcal{J}$ is nonzero. Then $S^*S \in \mathcal{J}$ is nonzero and $\tau(S^*S) > 0$ by the faithfulness. By the Claim, since unitary conjugations and convex combinations and limits of elements of $\mathcal{J}$ stay in $\mathcal{J}$, we get that $\tau(S^*S)\, I_{\mathcal{H}} \in \mathcal{J}$, and hence $\mathcal{J} = \mathcal{A}$, implying that $\mathcal{A}$ is simple.

So it remains to prove the Claim. Let $\mathcal{A}_0 = \lambda(\mathbb{C}\,\mathbb{F}_2)$. As $\mathcal{A}_0$ is dense in $\mathcal{A}$, it is enough to prove the Claim for $S \in \mathcal{A}_0$. We can rewrite (13.15) as

$$0 \in \overline{\mathrm{conv}}\{U(S - \tau(S)\, I_{\mathcal{H}})U^* : \ U \in \mathcal{A}_0 \ \text{ is unitary}\}, \tag{13.16}$$

and so it is enough to show (13.15) in the case where $\tau(S) = 0$, on top of the assumption that $S \in \mathcal{A}_0$. That is, since $S = S^*$,

$$S = \sum_{j=1}^{n} \alpha_j \lambda(g_j) + \bar{\alpha}\lambda(g_j^{-1}),$$

where $g_1, \ldots, g_n$ are distinct, and none equal to $e$ (this, from $\tau(S) = 0$). For each $g \in \mathbb{F}_2$ we denote by $\ell(g)$ the length of the word, and by $\pi_k(w)$ the $k^{\text{th}}$ letter of the word $w$. As before we write $a_1, a_2$ for the generators of $\mathbb{F}_2$. Let $r = a_1a_2$ and $s = a_1a_2^{-1}$. Put $j = 1 + \max\{\ell(g_1), \ldots, \ell(g_n)\}$. Let

$$z_k = r^k s^2 r^j, \qquad k \in \mathbb{N},$$

and

$$A_k = \{w \in \mathbb{F}_2 : \ \pi_1(r^{-k}w) = a_1, \ \text{ and } \pi_1(a_1^{-1}r^{-k}w) \neq a_2\}, \qquad k \in \mathbb{N}.$$

If $w \in A_{k_1} \cap A_{k_2}$ with $k_1 < k_2$, then $w = r^{k_1} a_1 x \cdots = r^{k_2} a_1 y \cdots$; so $a_1 x \cdots = r^{k_2 - k_1} a_1 y \cdots$, where $x, y \neq a_2$. Since $r = a_1 a_2$ and $k_2 - k_1 > 0$, the first two letters of the equality would give $a_1 x = a_1 a_2$, so $x = a_2$ which contradicts the definition of $A_{k_1}$. It follows that the $\{A_k\}$ are pairwise disjoint.

We let $P_k \in \mathcal{B}(\mathcal{H})$ be the multiplication operator by $1_{A_k}$; equivalently, $P_k$ is the orthogonal projection onto the subspace of words with support in $A_k$. The disjointness of the $A_k$ makes the family $\{P_k\}$ pairwise orthogonal. Take $\xi \in \mathcal{H}$, supported in $\mathbb{F}_2 \setminus A_k$, and $y \in \mathbb{F}_2 \setminus A_k$. We have $w = z_k g_h^{-1} z_k^{-1} y \in A_k$; indeed, $r^{-k} w = s^2 r^j g_h^{-1} z_k^{-1} y$ begins with $a_1$, and $a_1^{-1} r^{-k} = a_2^{-1} r^j g_h^{-1} z_k^{-1}$ does not begin with $a_2$. Thus $(1_{A_k^c} \xi)(z_k g_h^{-1} z_k^{-1} y) = 0$, which means that $(I_{\mathcal{H}} - P_k)\lambda(z_k g_h z_k^{-1})(I_{\mathcal{H}} - P_k) = 0$, for $h = 1, \ldots, n$. We can write this as $(I_{\mathcal{H}} - P_k)\lambda(z_k)\lambda(g_h)\lambda(z_k)^*(I_{\mathcal{H}} - P_k) = 0$, and taking adjoints and linear combinations we obtain

$$(I_{\mathcal{H}} - P_k)\lambda(z_k) S \lambda(z_k)^* (I_{\mathcal{H}} - P_k) = 0.$$

Then, using Exercise 10.5.14 in the second inequality,

$$\begin{aligned}\left|\left\langle \frac{1}{9}\sum_{k=1}^{9} \lambda(z_k) S \lambda(z_k)^* \xi, \xi \right\rangle\right| &\leq \frac{1}{9}\sum_{k=1}^{9} |\langle \lambda(z_k) S \lambda(z_k)^* \xi, \xi\rangle| \leq \frac{2}{9}\sum_{k=1}^{9} \|S\|\, \|P_k \xi\| \\ &\leq \frac{2}{9}\|S\|\sqrt{9}\Big(\sum_{k=1}^{9}\|P_k\xi\|^2\Big)^{1/2} \leq \frac{2}{3}\|S\|.\end{aligned}$$

As we are doing this for a selfadjoint operator and $\xi \in \mathcal{H}$ with $\|\xi\| \leq 1$ was arbitrary, we get that

$$\left\|\frac{1}{9}\sum_{k=1}^{9} \lambda(z_k) S \lambda(z_k)^*\right\| \leq \frac{2}{3}\|S\|.$$

If we now iterate the construction $m$ times, we get an element

$$S_m \in \operatorname{conv}\{\lambda(g) S \lambda(g)^* : \ g \in \mathbb{F}_2\} \qquad \text{with} \qquad \|S_m\| \leq \Big(\frac{2}{3}\Big)^m \|S\|,$$

and this shows that 0 is in the closure of the convex hull of the unitary orbit of $S$. □

*13.3.11. Remark* . The approximation of the trace by averages in (13.15), used by Powers to show the simplicity of $C^*_\lambda(\mathbb{F}_2)$, has been shown by Kennedy—more than 40 years later!—to precisely characterize the simplicity of $C^*_\lambda(G)$ for $G$ discrete [Ken20]. The equality (13.15) is particular case of the Dixmier Property, see Section 14.5.

***13.3.12. Proposition.***

*Let $\mathcal{A} = C^*_\lambda(\mathbb{F}_2)$, and $\pi : \mathcal{A} \otimes \mathcal{A} \to \mathcal{B}(\mathcal{H}_{\mathbb{F}_2})$ the representation induced by $\pi(x \otimes y) = xJyJ$. Then $\pi$ is irreducible.*

*Proof.* A typical application of Proposition 13.1.4 shows that $\pi$ is well-defined. Indeed, if $\sum_k x_k \otimes y_k = 0$, then there exist coefficients $\{c_{kj}\}$ such that

$$\sum_k c_{kj} x_j = 0, \qquad \sum_j c_{kj} y_j = y_k.$$

Then

$$\sum_k x_k J y_k J = \sum_{k,j} c_{kj} x_k J y_j J = \sum_j \Big(\sum_k c_{kj} x_j\Big) J y_j J = 0.$$

The image of $\pi$ is $\mathrm{alg}\{C^*_\lambda(\mathbb{F}_2) \cup C^*_\rho(\mathbb{F}_2)\} \subset \mathcal{B}(\mathcal{H}_{\mathbb{F}_2})$. Let $T \in \pi(\mathcal{A} \otimes \mathcal{A})'$. This means that $T\lambda(g) = \lambda(g)T$ and $T\rho(g) = \rho(g)T$ for all $g \in \mathbb{F}_2$. For any $g \in \mathbb{F}_2$ we have

$$\lambda(g)\rho(g)T\delta_e = T\lambda(g)\rho(g)\delta_e = T\delta_{gg^{-1}} = T\delta_e.$$

Write $T\delta_e = \displaystyle\sum_g \alpha_g \delta_g$. Then, for any $h \in \mathbb{F}_2$,

$$\sum_g \alpha_g \delta_g = T\delta_e = \sum_g \alpha_g \lambda(h)\rho(h)\delta_g = \sum_g \alpha_g \delta_{hgh^{-1}} = \sum_g \alpha_{h^{-1}gh} \delta_g.$$

As the coefficients are unique, $\alpha_{h^{-1}gh} = \alpha_g$ for all $g, h \in \mathbb{F}_2$. As we are in a free group, the conjugacy class of $g \neq e$ is infinite, and so

$$\|T\delta_e\| \geq \sum_h |\alpha_{h^{-1}gh}|^2 = \sum_h |\alpha_g|^2,$$

which is only possible if $\alpha_g = 0$. It follows that $T\delta_e = \alpha\delta_e$, and then

$$T\delta_g = T\lambda(g)\delta_e = \lambda(g)T\delta_e = \alpha\lambda(g)\delta_e = \alpha\delta_g$$

for all $g$, showing that $T = \alpha I_{\mathcal{H}_{\mathbb{F}_2}}$. We have shown that $\pi(\mathcal{A} \otimes \mathcal{A})' = \mathbb{C}\, I_{\mathcal{H}_{\mathbb{F}_2}}$, and so $\pi$ is irreducible. ☐

**13.3.3. Some nontrivial von Neumann algebras.** The ideas in this section allow us to construct von Neumann algebras that look nothing like the examples provided in Examples 12.3.7.

*13.3.13. Example*. Let $G$ be a countable group. We require it to be **icc**, which stands for *infinite conjugacy classes*. By this we mean that for any $g \in G$ other than the unit, the conjugacy class $C(g) = \{hgh^{-1} : h \in G\}$ is infinite. These groups are common: for example we can take $G = \mathbb{S}_\infty$ (permutations of $\mathbb{N}$ that only alter finite many elements) or the free groups $\mathbb{F}_n$; and there are many other examples. Consider the left regular representation of $G$, so we get a group of unitaries $\lambda(G) \subset \mathcal{B}(\mathcal{H}_G)$. Then $L(G) = \lambda(G)''$ is a von Neumann algebra. It is actually the sot-closure of the left reduced C*-algebra of $G$. Similarly, $R(G) = \rho(G)''$ is a von Neumann algebra, the sot-closure of the right reduced $C^*$-algebra of $G$.

We considered convolutions a few times already; now we will define it in the context of functions on groups. Concretely, if $c, d : G \to \mathbb{C}$ are functions, the **left convolution** is the function $c * d : G \to \mathbb{C}$ given by

$$(c * d)(g) = \sum_{h \in G} c(h) d(h^{-1} g).$$

We can analogously define the right convolution. Of course in general $G$ will be infinite and the series is not guaranteed to converge. But if $c, d \in \ell^2(G)$, then Cauchy–Schwarz gives us that $\|c * d\|_\infty \leq \|c\|_2 \|d\|_2$. Which means that the convolution may fall outside of $\ell^2(G)$, but at least the series exists and the values are uniformly bounded.

Suppose that $T \in R(G)'$. Let $\xi = T\delta_e$ and let $c \in \ell^2(G)$ with $T\delta_e = \sum_h c_h \delta_h$. For any $g \in G$,

$$T\delta_g = T\rho(g^{-1})\delta_e = \rho(g^{-1}) T\delta_e = \sum_{h \in G} c_h \delta_{hg} = \sum_{h \in G} c_h \lambda(h)\, \delta_g.$$

By linearity this extends to $T\eta = \Big(\sum_h c_h \lambda(h)\Big)\eta$ for all $\eta \in \operatorname{span}\{\delta_g : g \in G\}$, with the series converging in the sot (as we have norm convergence at the level of vectors in the equality above). More generally, given $\eta = \sum_{g \in G} \alpha_g \delta_g \in \ell^2(G)$ and using the continuity of $T$,

$$\begin{aligned} T\eta &= \sum_{g \in G} \alpha_g T\delta_g = \sum_{g \in G} \sum_{h \in G} c_h \alpha_g \delta_{hg} = \sum_{g \in G} \sum_{h \in G} c_h \alpha_{h^{-1}g} \delta_g \\ &= c * \eta \end{aligned} \tag{13.17}$$

(depending on point of view, the above computation might look shaky with the double series, even though there is not exchange; we refer to Exercise 13.3.9 for an alternative approach). So $c$ has the additional property that convolution against any element of $\ell^2(G)$ is again in $\ell^2(G)$. Recall that since $G$ is non-abelian, convolution is not commutative.

If we consider, for $F \subset G$ finite, the partial sums $T_F = \sum_{g \in F} c_h \lambda(h)$, we have that $T_F \eta = (P_F c) * \eta$, where $P_F \in \mathcal{B}(\ell^2(G))$ is the projection $P_F \xi = 1_F \xi$. By Exercise 13.3.6,

$$\|T_F \eta\|_2 = \|(P_F c) * \eta\|_2 = \|c * (P_{F^{-1}g} \eta)\|_2 = \|T P_{F^{-1}g} \eta\|_2 \leq \|T\|\, \|\eta\|_2.$$

Then by Exercise 12.1.6 we have that $T = \sum_h c_h \lambda(h)$, with the series converging sot. In particular $T \in L(G)$. So $R(G)' \subset L(G)$. The reverse inclusion is trivial, from $\lambda(g)\rho(h) = \rho(h)\lambda(g)$, and therefore $L(G) = R(G)'$. Taking commutants we also get $L(G)' = R(G)$. The operators in $R(G)$ can be seen as right convolutions via analogous ideas.

The algebras $L(G)$ and $R(G)$ are non-commutative. In fact, they are factors. To see this, suppose that $Z \in \mathcal{Z}(L(G))$. We have $Z = \sum_{h \in G} z_h \lambda(h)$. Fix any nontrivial $g \in G$. Since $Z$ is central, $Z = \lambda(g) Z \lambda(g)^{-1}$. By Exercise 13.3.8 the coefficients in the series representation are unique. Since

$$\sum_{h \in G} z_h \lambda(h) = Z = \lambda(g) Z \lambda(g^{-1}) = \sum_{h \in G} z_h \lambda(ghg^{-1}) = \sum_{h \in G} z_{g^{-1}hg} \lambda(h),$$

it follows that $z_h = z_{ghg^{-1}}$ for all $g, h \in G$. But this is a problem, because we are assuming that our group is icc, so the conjugacy class of each nontrivial $h$ is infinite. This means that the sequence $z$ contains infinitely many copies of the number $z_h$ for each $h \neq e$. As $z \in \ell^2(G)$, this means that $z_h = 0$ for all $h \neq e$. That is, $Z = z_e \lambda(e) = z_e \, I_{\ell^2(G)}$. So $\mathcal{Z}(L(G)) = \mathbb{C} \, I_{\ell^2(G)}$. And we also get $\mathcal{Z}(R(G)) = R(G) \cap R(G)' = L(G)' \cap L(G) = \mathcal{Z}(L(G)) = \mathbb{C} \, I_{\ell^2(G)}$.

We already know infinite-dimensional von Neumann algebras with trivial centre, namely $\mathcal{B}(\mathcal{H})$ for infinite-dimensional $\mathcal{H}$. But here is a way to see that these new von Neumann algebras are nothing like that. Consider the canonical trace $\tau : L(G) \to \mathbb{C}$ given by $\tau(T) = \langle T\delta_e, \delta_e \rangle$. We saw in Proposition 13.3.6 that this is bounded, and on $C^*_\lambda(G)$ it is tracial, meaning that $\tau(ST) = \tau(TS)$. The elements of $L(G)$ are wot limits of elements in $C^*_\lambda(G)$, so for $S, T \in L(G)$ and nets $\{T_j\}, \{S_k\} \subset C^*_\lambda(G)$ with $S_k \xrightarrow{\text{sot}} S$ and $T_k \xrightarrow{\text{sot}} T$,

$$\begin{aligned}\tau(ST) = \langle T\delta_e, S^*\delta_e \rangle &= \lim_k \lim_j \langle T_j \delta_e, S_k^* \delta_e \rangle = \lim_k \lim_j \tau(S_k T_j) = \lim_k \lim_j \tau(T_j S_k) \\ &= \lim_k \lim_j \langle S_k \delta_e, T_k^* \delta_e \rangle = \tau(TS).\end{aligned}$$

So $L(G)$ has a tracial state. This makes it radically different from $\mathcal{B}(\mathcal{H})$. We will see as we learn more about von Neumann algebras that the differences pile up. For instance $L(G)$ has no minimal projections, and it is simple as a C$^*$-algebra (Proposition 14.7.2). It is an example of a II$_1$-factor (cfr. Remark 14.3.8).

A rather explicit description of $L(G)$ is as the set of "operators with Toeplitz matrices", or "constant along diagonals". By this we mean (Exercise 13.3.11)

$$L(G) = \{T \in \mathcal{B}(\ell^2(G)) : \; T_{g,hg} = T_{k,hk}, \; g, k, h \in G\},$$

where $T_{g,h} = \langle T\delta_h, \delta_g \rangle$. The elements $\{T_{g,hg}\}_{g \in G}$, for fixed $h \in G$, can be though of as one diagonal of the matrix of $T$ in the canonical basis (think of the usual matrix, where a diagonal is $T_{k,k+r}$).

**13.3.4. Exercises.**

(13.3.1) Prove Proposition 13.3.2.

(13.3.2) Prove Proposition 13.3.4.

(13.3.3) Let $x \in C^*_\lambda(G)$ and $y \in C^*_\rho(G)$. Show that $xy = yx$.

(13.3.4) Let $J$ be given by $(J\xi)(g) = \xi(g^{-1})$. Show that $J \in \mathcal{B}(\mathcal{H}_G)$ is a unitary and $J\lambda(g)J = \rho(g)$ for all $g \in G$.

(13.3.5) Let $G$ be a discrete group, $c \in \ell^2(G)$ with the property that $c*\eta \in \ell^2(G)$ for all $\eta \in \ell^2(G)$, and $T : \ell^2(G) \to \ell^2(G)$ the operator $T\eta = c * \eta$. Use the Closed Graph Theorem to show that $T$ is bounded.

(13.3.6) Let $c, \eta \in \ell^2(G)$, $F \subset G$ and $P_F \in \mathcal{B}(\ell^2(G))$ the projection $(P_F\xi)(g) = 1_F(g)\xi(g)$. Show that $(P_F c) * \eta = c * (P_{F^{-1}g}\eta)$.

(13.3.7) For each $g \in G$ let $P_g$ be the orthogonal projection onto $\mathbb{C}\,\delta_g$. Show that

$$\lambda(g)P_e\lambda(g)^* = P_g, \qquad \text{and} \qquad \rho(g)\,P_e\,\rho(g)^* = P_{g^{-1}}.$$

(13.3.8) Show that the tracial state $\tau$ is faithful on $L(G)$.

(13.3.9) Write an alternative proof of (13.17) by writing each coordinate of $T\eta$.

(13.3.10) Let $G$ be a discrete group. The **full $\mathbf{C}^*$-algebra** of $G$ is the completion $C^*(G)$ of $\mathbb{C}G$ via the norm

$$\Big\|\sum_{g\in G} a_g\, g\Big\| = \sup\Big\{\Big\|\sigma\Big(\sum_{g\in G} a_g\, g\Big)\Big\| :\ \sigma \in R_G\Big\},$$

where $R_G$ is the set of $*$-representations $\sigma : \mathbb{C}G \to \ell^2(S_n)$ where each $S_n$ is a set of cardinality $n$ for each $n \leq |G|$ (these convoluted choice guarantees that the set of representations is actually a set). Show that $C^*_\lambda(G)$ and $C^*_\rho(G)$ are quotients of $C^*(G)$.

(13.3.11) Let $G$ be a discrete group and $T \in \mathcal{B}(\ell^2(G))$. Show that $T \in L(G)$ if and only if $T$ is "Toeplitz", in the sense that diagonals are constant, meaning that for $g, k, h \in G$,

$$\langle T\delta_{hg}, \delta_g\rangle = T_{g,hg} = T_{k,hk} = \langle T\delta_{hk}, \delta_h\rangle. \tag{13.18}$$

## 13.4. Topological Tensor Products

There is a notational issue with tensor products, which is that one needs to distinguish algebraic tensor products from those obtained as a completion. We will save the plain symbol $\otimes$ to denote the algebraic tensor product, and qualify it in different ways for other tensor products.

Here we will also list the several subsections for easy access:

### 13.4.1. Tensor Products of Hilbert Spaces.

**13.4.1. Definition.**

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. The* ***Hilbert space tensor product*** $\mathcal{H} \bar{\otimes} \mathcal{K}$ *is the Hilbert space obtained as the completion of $\mathcal{H} \otimes \mathcal{K}$ with respect to the inner product*

$$\left\langle \sum_j \xi_j \otimes \eta_j, \sum_k \nu_k \otimes \zeta_k \right\rangle = \sum_{k,j} \langle \xi_j, \nu_k \rangle \langle \eta_j, \zeta_k \rangle. \tag{13.19}$$

One needs to be careful here, for it is not a priori entirely clear that (13.19) actually defines an inner product. Let us address that first, and let us get a glimpse of how messy tensor products can be.

***13.4.2. Proposition.***

*The formula (13.19) does define an inner product on $\mathcal{H} \otimes \mathcal{K}$, making it a pre-Hilbert space.*

*Proof.* Note that all sums in the proof are finite. The sequilinearity is straightforward to check. The issue is whether the inner product is well-defined and positive-definite. For this, suppose that

$$\sum_{j=1}^n \xi_j \otimes \eta_j = \sum_{r=1}^m \xi_r' \otimes \eta_r'.$$

By relabelling $\xi_{n+j} = -\xi_j'$ and $\eta_{n+j} = \eta_j'$, we may write the above as

$$\sum_{j=1}^{n+m} \xi_j \otimes \eta_j = 0.$$

Then Proposition 13.1.4 gives us coefficients $\{c_{kj}\}$ such that

$$\sum_{k=1}^{n+m} c_{kj}\xi_k = 0, \qquad \sum_{j=1}^{n+m} c_{kj}\eta_j = \eta_k.$$

Now

$$\begin{aligned}
\left\langle \sum_{j=1}^{n} \xi_j \otimes \eta_j, \sum_{k=1}^{p} \nu_k \otimes \zeta_k \right\rangle &= \sum_{j=1}^{n}\sum_{k=1}^{p} \langle \xi_j, \nu_k\rangle\langle \eta_j, \zeta_k\rangle \\
&= \sum_{j=1}^{n}\sum_{k=1}^{p}\sum_{s=1}^{n+m} c_{js}\langle \xi_j, \nu_k\rangle\langle \eta_s, \zeta_k\rangle \\
&= \sum_{s=1}^{n+m}\sum_{k=1}^{p}\left\langle \sum_{j=1}^{n} c_{js}\xi_j, \nu_k\right\rangle\langle \eta_s, \zeta_k\rangle \\
&= \sum_{s=1}^{n+m}\sum_{k=1}^{p}\left\langle -\sum_{j=n+1}^{n+m} c_{js}\xi_j, \nu_k\right\rangle\langle \eta_s, \zeta_k\rangle \\
&= \sum_{j=n+1}^{n+m}\sum_{k=1}^{p}\langle -\xi_j, \nu_k\rangle\left\langle \sum_{s=1}^{n+m} c_{js}\eta_s, \zeta_k\right\rangle \\
&= \sum_{j=n+1}^{n+m}\sum_{k=1}^{p}\langle -\xi_j, \nu_k\rangle\langle \eta_j, \zeta_k\rangle \\
&= \sum_{j=1}^{m}\sum_{k=1}^{p}\langle \xi_j', \nu_k\rangle\langle \eta_j', \zeta_k\rangle \\
&= \left\langle \sum_{j=1}^{m} \xi_j' \otimes \eta_j', \sum_{k=1}^{p} \nu_k \otimes \zeta_k \right\rangle.
\end{aligned}$$

Repeating the argument on the second entry of the inner product allows us to show that if $\sum_j \xi_j \otimes \eta_j = \sum_r \xi_r' \otimes \eta_r'$ and $\sum_k \nu_k \otimes \zeta_k = \sum_s \nu_s' \otimes \zeta_s'$, then

$$\sum_{k,j}\langle \xi_j, \nu_k\rangle\langle \eta_j, \zeta_k\rangle = \sum_{r,s}\langle \xi_r', \nu_s'\rangle\langle \eta_r', \zeta_s'\rangle,$$

so the form is well-defined. Finally, if

$$\left\langle \sum_j \xi_j \otimes \eta_j, \sum_j \xi_j \otimes \eta_j \right\rangle = 0, \qquad \text{then} \qquad \sum_{k,j}\langle \xi_j, \xi_k\rangle\langle \eta_j, \eta_k\rangle = 0.$$

Because the inner product does not depend on the representative, we may choose $\{\eta_j\}$ to be orthonormal. Then the equality becomes $\sum_j \|\xi_j\|^2 = 0$, so $\xi_j = 0$ for all $j$, and then $\sum_j \xi_j \otimes \eta_j = 0$. This business of choosing the $\eta_k$ to be orthonormal also shows that the form is positive, so it is positive-definite. $\square$

If $\{\xi_j\}$ is an orthonormal basis for $\mathcal{H}$ and $\{\eta_k\}$ is an orthonormal basis for $\mathcal{K}$, then $\{\xi_j \otimes \eta_k\}$ is an orthonormal basis for $\mathcal{H} \overline{\otimes} \mathcal{K}$ (Exercise 13.4.1). This highlights the difference between the direct sum and the tensor product, as the former has dimension $\dim \mathcal{H} + \dim \mathcal{K}$, while the latter has dimension $\dim \mathcal{H} \dim \mathcal{K}$.

Though we just established how to get orthonormal bases on the tensor product, sometimes a different form to express a vector will be desirable. We have

***13.4.3. Lemma.***

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $\xi \in \mathcal{H} \overline{\otimes} \mathcal{K}$. Given an orthonormal basis $\{\eta_k\}$ of $\mathcal{K}$, there exist unique vectors $\{\zeta_k\} \subset \mathcal{H}$ such that*

$$\xi = \sum_k \zeta_k \otimes \eta_k$$

*and $\sum_k \|\zeta_k\|^2 = \|\xi\|^2$.*

*Proof.* Fix an orthonormal basis $\{\xi_j\}$ for $\mathcal{H}$. By Exercise 13.4.1 we can write $\xi = \sum_{k,j} \alpha_{kj}\, \xi_j \otimes \eta_k$ for coefficients $\alpha_{kj}$ that satisfy $\sum_{k,j} |\alpha_{kj}|^2 = \|\xi\|^2$. Let

$$\zeta_k = \sum_j \alpha_{kj} \xi_j.$$

This series converges in $\mathcal{H}$ for each $k$, and in fact $\sum_k \|\zeta_k\|^2 = \sum_{k,j} |\alpha_{kj}|^2 = \|\xi\|^2$. Now, for $F$ finite

$$\begin{aligned} \Big\| \xi - \sum_{k \in F} \zeta_k \otimes \eta_k \Big\|^2 &= \Big\| \sum_{k \notin F} \sum_j \alpha_{kj}\, \xi_j \otimes \eta_k \Big\|^2 \\ &= \sum_{k \notin F} \sum_j |\alpha_{kj}|^2 = \sum_{k \notin F} \|\zeta_k\|^2 \xrightarrow[F]{} 0. \end{aligned}$$

For the uniqueness, suppose that $\xi = \sum_k \zeta_k \otimes \eta_k = \sum_k \zeta_k' \otimes \eta_k$. For any $\nu \in \mathcal{H}$ and fixed $k$ we have

$$\langle \zeta_k, \nu \rangle = \langle \xi, \nu \otimes \eta_k \rangle = \langle \zeta_k', \nu \rangle.$$

Hence $\zeta_k' = \zeta_k$. $\square$

The norm we defined on $\mathcal{H} \otimes \mathcal{K}$ is a **cross norm**, in the sense that $\|\xi \otimes \eta\| = \|\xi\|\, \|\eta\|$. This follows directly from the definition. We remark that though most tensor product norms we will see will be cross norms, they do not have to be necessarily.

We also mention that since $\mathcal{H} \otimes \mathcal{K}$ is canonically isomorphic to $\mathcal{K} \otimes \mathcal{H}$ (Exercise 13.1.4), everything we do on the left of the tensor product we can do on the right, and vice versa.

If $\{\xi_j\}_{j\in J}$ and $\{\eta_k\}_{k\in K}$ are orthonormal bases of $\mathcal{H}$ and $\mathcal{K}$ respectively, there are canonical isomorphisms

$$\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K} \simeq \bigoplus_{j\in J} \mathcal{K} \simeq \bigoplus_{k\in K} \mathcal{H} \tag{13.20}$$

(Exercise 13.4.2). Tensor products have already appeared, under the disguise of (13.20), in Corollary 12.6.16.

**13.4.2. Tensor Products of Operators.** Given Hilbert spaces $\mathcal{H}, \mathcal{K}$ and $T \in \mathcal{B}(\mathcal{H}), S \in \mathcal{B}(\mathcal{K})$, it looks natural to define a linear operator by $(T \otimes S)(\xi \otimes \eta) = T\xi \otimes S\eta$ and extend by linearity. This, in fact, we already did in Corollary 13.1.8. The problem is that now we have a Hilbert space $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$ and we want $T \otimes S \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$.

***13.4.4. Proposition.***

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $T \in \mathcal{B}(\mathcal{H})$, $S \in \mathcal{B}(\mathcal{K})$. Then there exists a unique linear operator $T \otimes S \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ such that $(T \otimes S)(\xi \otimes \eta) = T\xi \otimes S\eta$. We also have $\|T \otimes S\| = \|T\|\,\|S\|$.*

*Proof.* Consider first the operator $T \otimes I_{\mathcal{K}} : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$. If we fix $\xi \in \mathcal{H} \otimes \mathcal{K}$, there exists an orthonormal basis $\{\eta_k\}$ for $\mathcal{K}$ such that $\xi = \sum_{j=1}^n \xi_j \otimes \eta_j$ (this follows from Exercise 13.1.2). Then

$$\begin{aligned} \big\|(T \otimes I_{\mathcal{K}})\xi\big\|^2 &= \Big\|\sum_{j=1}^n T\xi_j \otimes \eta_j\Big\|^2 = \sum_{j=1}^n \|T\xi_j\|^2 \\ &\leq \|T\|^2 \sum_{j=1}^n \|\xi_j\|^2 = \|T\|^2 \Big\|\sum_{j=1}^n \xi_j \otimes \eta_j\Big\|^2 = \|T\|^2\,\|\xi\|. \end{aligned}$$

This shows that $T \otimes I_{\mathcal{K}}$ is bounded on $\mathcal{H} \otimes \mathcal{K}$ and hence by Proposition 6.1.9 it extends uniquely to $T \otimes I_{\mathcal{K}} \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$, with $\|T \otimes I_{\mathcal{K}}\| \leq \|T\|$. We show similarly that there is a unique $I_{\mathcal{H}} \otimes S \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ that extends $I_{\mathcal{K}} \otimes S$, and $\|I_{\mathcal{H}} \otimes S\| \leq \|S\|$. With this, we define $T \otimes S = (T \otimes I_{\mathcal{K}})(I_{\mathcal{H}} \otimes S) \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$. Given $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$ we have

$$(T \otimes S)(\xi \otimes \eta) = (T \otimes I_{\mathcal{K}})(I_{\mathcal{H}} \otimes S)(\xi \otimes \eta) = (T \otimes I_{\mathcal{K}})(\xi \otimes S\eta) = T\xi \otimes S\eta.$$

As for the norm, from the construction we get $\|T \otimes S\| \leq \|T\|\,\|S\|$. Given $\varepsilon > 0$ there exist unit vectors $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$ with $\|T\xi\| \geq (1-\varepsilon)\|T\|$ and $\|S\eta\| \geq (1-\varepsilon)\|S\|$. Then $\|\xi \otimes \eta\| = \|\xi\|\,\|\eta\| = 1$, and

$$\|(T \otimes S)(\xi \otimes \eta)\| = \|T\xi \otimes S\eta\| = \|T\xi\|\,\|S\eta\| \geq (1-\varepsilon)^2\|T\|\,\|S\|.$$

As $\varepsilon$ was arbitrary this shows the reverse inequality and hence $\|T \otimes S\| = \|T\| \, \|S\|$. $\square$

Tensor products of operators satisfy the same arithmetic properties than the tensor product of vectors, as in Proposition 13.1.2 (Exercise 13.4.3). A direct computation shows that

$$(T \otimes S)^* = T^* \otimes S^*.$$

*13.4.5. Remark*. There is an issue to discuss regarding Proposition 13.4.4 and the paragraph preceding it. Given $T \in \mathcal{B}(\mathcal{H})$ and $S \in \mathcal{B}(\mathcal{K})$, we may consider the element $T \otimes S$ in the algebra $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$. This does not immediately make it clear how to see $T \otimes S$ as an operator. It can be done, though, for we may identify (via Exercise 13.1.5) $\mathcal{B}(\mathcal{H})$ with $\mathcal{B}(\mathcal{H}) \otimes \mathbb{C} I_{\mathcal{K}}$ and $\mathcal{B}(\mathcal{K})$ with $\mathbb{C} I_{\mathcal{H}} \otimes \mathcal{B}(\mathcal{K})$, so $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ is naturally a subalgebra of $\mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$, via Proposition 13.4.4. In general—that is, in infinite-dimension—it is a proper (not dense) subalgebra (see Exercise 13.4.11).

**13.4.3. Tensor Products of von Neumann Algebras.** Since we consider von Neumann algebras as concrete subalgebras of operators on a Hilbert space, there is only one canonical way of defining the tensor product. Namely, if $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$, then $\mathcal{M} \otimes \mathcal{N}$ sits naturally in $\mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$. So

**13.4.6. Definition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be von Neumann algebras. The von Neumann algebra $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N} \subset \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ is defined to be $\overline{\mathcal{M} \otimes \mathcal{N}}^{\text{sot}}$.*

Some elementary and useful relations are

- if $\dim \mathcal{K} = n$, then $\mathcal{M} \otimes \mathcal{B}(\mathcal{K}) \simeq M_n(\mathcal{M})$ (Exercise 13.1.5);
- $(\mathcal{M} \otimes I_{\mathcal{K}})' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$;
- $(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}))' = \mathcal{M}' \otimes I_{\mathcal{K}}$

(Exercise 13.4.27). The notation $\mathcal{M} \otimes I_{\mathcal{K}}$ is a bit awkward in the sense that it should properly be $\mathcal{M} \otimes \mathbb{C}\, I_{\mathcal{K}}$; but since scalars "float" inside a tensor product, every element in such algebra is of the form $X \otimes I_{\mathcal{K}}$ with $X \in \mathcal{M}$.

The main result in this context is the fact that the commutant of the product is the product of the commutants:

$$(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}'. \tag{13.21}$$

We will prove this in Theorem 14.4.3.

The tensor product $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N}$ of two von Neumann algebras has $\mathcal{M} \otimes I_{\mathcal{N}}$ and $I_{\mathcal{M}} \otimes \mathcal{N}$ as canonical subalgebras. If $\varphi \in \mathcal{M}_*$ and $\psi \in \mathcal{N}_*$, by Proposition 12.6.3 there exist trace-class operators $S_\varphi \in \mathcal{T}(\mathcal{H})$ and $S_\psi \in \mathcal{T}(\mathcal{K})$ such that $\varphi =$

$\operatorname{Tr}(S_\varphi\cdot)$ and $\psi=\operatorname{Tr}(S_\psi\cdot)$. Then we can form $\varphi\overline{\otimes}\psi:\mathcal{M}\overline{\otimes}\mathcal{N}\to\mathbb{C}$ by $(\varphi\overline{\otimes}\psi)(X)=\operatorname{Tr}((S_\varphi\otimes S_\psi)X)$. As the operator $S_\varphi\otimes S_\psi$ is trace-class (Exercise 13.4.7), the functional $\varphi\overline{\otimes}\psi$ is normal. Now fix $\varphi\in\mathcal{M}_*$ and $\tilde X\in\mathcal{M}\overline{\otimes}\mathcal{N}$ and consider on $\mathcal{N}_*$ the map $\psi\longmapsto(\varphi\otimes\psi)(\tilde X)$. As this is bounded and $(\mathcal{N}_*)^*=\mathcal{N}$ (Theorem 12.7.2), there exists a unique $\mathcal{E}_{\mathcal{N}}^{\varphi}(X)\in\mathcal{N}$ such that

$$\psi(\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X))=(\varphi\otimes\psi)(\tilde X),\qquad \psi\in\mathcal{N}_*. \tag{13.22}$$

That is, we have a map $\mathcal{E}_{\mathcal{N}}^{\varphi}:\mathcal{M}\,\overline{\otimes}\,\mathcal{N}\to\mathcal{N}$. Since the equality (13.22) is linear on $\tilde X$, the uniqueness guarantees that $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is linear. Also

$$\|\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X)\|=\sup\{|\psi(\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X))|:\ \varphi\in\mathcal{N}_*,\ \|\psi\|=1\}\le\|\varphi\|\,\|\tilde X\|.$$

For $X\in\mathcal{M}$ and $Y\in\mathcal{N}$, (13.22) gives (again by the uniqueness)

$$\mathcal{E}_{\mathcal{N}}^{\varphi}(X\otimes Y)=\varphi(X)\,Y.$$

So if we take $\varphi$ to be a normal state on $\mathcal{M}$ and we identify $\mathcal{N}$ with $I_{\mathcal{M}}\otimes\mathcal{N}$, Proposition 13.2.68 gives us that $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is a conditional expectation onto $I_{\mathcal{M}}\otimes\mathcal{N}$. We can see the conditional expectation $\mathcal{E}_{\mathcal{N}}^{\varphi}$ as the map $\varphi\,\overline{\otimes}\,\mathrm{id}_{\mathcal{N}}$. It is often called a **slice map**, and also a **Fubini map**. In the context of Quantum Information and when $\varphi$ is a trace, the map $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is called a **partial trace**.

***13.4.7. Proposition.***

*Let $\mathcal{M},\mathcal{N}$ be von Neumann algebras and $\varphi\in\mathcal{M}_*^+$. Then $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is normal. If $\varphi$ is faithful then $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is faithful.*

*Proof.* For the normality, by definition we have $\psi\circ\mathcal{E}_{\mathcal{N}}^{\varphi}=\varphi\otimes\psi$, and the latter is normal as discussed in the paragraph after (13.21). So if $\tilde X_j\xrightarrow{\sigma\text{-weak}}0$ then $\psi(\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X_j))\to0$ for all $\psi\in\mathcal{N}_*$ (using Corollary 12.6.5), and this means by definition that $\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X_j)\xrightarrow{\sigma\text{-weak}}0$. So $\mathcal{E}_{\mathcal{N}}^{\varphi}$ is normal.

Now suppose that $\mathcal{E}_{\mathcal{N}}^{\varphi}(\tilde X^*\tilde X)=0$. By (13.22), $(\varphi\otimes\psi)(\tilde X^*\tilde X)=0$ for all $\psi\in\mathcal{N}_*$. We can see this as

$$(\varphi\otimes\mathrm{id}_{\mathbb{C}})(\mathrm{id}_{\mathcal{M}}\otimes\psi)(\tilde X^*\tilde X)=0. \tag{13.23}$$

The map $\varphi\otimes\mathrm{id}_{\mathbb{C}}$ acts on $\mathcal{M}\otimes\mathbb{C}$ via $(\varphi\otimes\mathrm{id}_{\mathbb{C}})(X\otimes1)=\varphi(X)$ so it can be naturally identified with $\varphi$; hence it is faithful. Therefore $(\mathrm{id}_{\mathcal{M}}\otimes\psi)(\tilde X^*\tilde X)=0$ by the faithfulness of $\varphi$ and the positivity of $\mathcal{E}_{\mathcal{M}}^{\psi}$. For any $\gamma\in S(\mathcal{M})$, we now have

$$(\gamma\otimes\psi)(X^*X)=(\gamma\otimes\mathrm{id}_{\mathbb{C}})(\mathrm{id}_{\mathcal{M}}\otimes\psi)(\tilde X^*\tilde X)=0.$$

With the usual assumption that $\mathcal{M}\subset\mathcal{B}(\mathcal{H})$ and $\mathcal{N}\subset\mathcal{B}(\mathcal{K})$ we have that $\mathcal{M}\,\overline{\otimes}\,\mathcal{N}\subset\mathcal{B}(\mathcal{H}\,\overline{\otimes}\,\mathcal{K})$. If in particular we take $\gamma$ and $\psi$ to be vector states, we get that $\langle\tilde X^*\tilde X\,(\xi\otimes\eta),\xi\otimes\eta\rangle=0$ for all $\xi\in\mathcal{H}$ and $\eta\in\mathcal{K}$. As $\mathcal{H}\,\overline{\otimes}\,\mathcal{K}$ consists of limits of linear combinations of vectors of the form $\xi\otimes\eta$, it follows that $\tilde X^*\tilde X=0$, and hence $\tilde X=0$, making $\mathcal{E}_{\mathcal{N}}^{\varphi}$ faithful. $\square$

**13.4.4. Tensor Products of C*-algebras.** Given C*-algebras $\mathcal{A}, \mathcal{B}$, as they are vector spaces we already have $\mathcal{A} \otimes \mathcal{B}$ as a vector space. We can put a product on $\mathcal{A} \otimes \mathcal{B}$ by letting $(a_1 \otimes b_1)(a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2$ and defining products of sums of elementary tensors in terms of these. Similarly, we can define an involution by $(a \otimes b)^* = a^* \otimes b^*$ and again letting the involution of a sum of elementary tensors to be the sum of the involutions of the elementary tensors. One needs some care, though, because as we have seen the same element in the tensor product can be expressed as different sums of elementary tensors. So we need an argument similar to the one used in the proof of Proposition 13.4.2, that we leave as an exercise (Exercise 13.4.14, or Exercise 13.4.15 for an alternative way). With those issues out of the way, we now have $\mathcal{A} \otimes \mathcal{B}$ as a $*$-algebra. But to have a C*-algebra we need a C*-norm, and for the algebra to be complete in that norm. We will see that such a norm always exists; but it turns out that, in general, it is not unique.

In order to discuss norms on tensor products of C*-algebras, information about representations will be needed.

***13.4.8. Proposition. (Product of Representations)***

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{K})$ representations. Then there exists a representation $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}} : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ such that $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(a \otimes b) = \pi_{\mathcal{A}}(a) \otimes \pi_{\mathcal{B}}(b)$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. If $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ are faithful, then $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$ is faithful.*

*Proof.* We can use Corollary 13.1.8 and Remark 13.4.5 to obtain a homomorphism $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}} : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K}) \subset \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ that satisfies $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(a \otimes b) = \pi_{\mathcal{A}}(a) \otimes \pi_{\mathcal{B}}(b)$. We also have

$$\big((\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(a \otimes b)\big)^* = \pi_{\mathcal{A}}(a)^* \otimes \pi_{\mathcal{B}}(b)^* = \pi_{\mathcal{A}}(a^*) \otimes \pi_{\mathcal{B}}(b^*) = (\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(a^* \otimes b^*),$$

so $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$ is a representation.

Now suppose that $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ are faithful. If $(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(x) = 0$, writing $x = \sum_{k=1}^{m} a_k \otimes b_k$ we have

$$0 = \sum_{k=1}^{m} \pi_{\mathcal{A}}(a_k) \otimes \pi_{\mathcal{B}}(b_k).$$

By Proposition 13.1.4 there exist numbers $c_{kj}$ such that

$$\sum_{k=1}^{m} c_{kj} \pi_{\mathcal{A}}(a_k) = 0 \quad \text{for all } j, \text{ and } \sum_{j=1}^{m} c_{kj} \pi_{\mathcal{B}}(b_j) = \pi_{\mathcal{B}}(b_k) \quad \text{for all } k.$$

As $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ are linear and faithful, we get

$$\sum_{k=1}^{m} c_{kj} a_k = 0 \text{ for all } j, \text{ and } \sum_{j=1}^{m} c_{kj} b_j = b_k \text{ for all } k,$$

and then $x = 0$ by Proposition 13.1.4. $\square$

Let us keep in mind that $\mathcal{A} \otimes \mathcal{B}$ is not a C*-algebra (unless one of them is finite-dimensional) and that it is not clear that $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$ is bounded; see Exercise 11.4.6 for an example of how things can go wrong in a situation like this, even without tensor products. As a kind of converse to Proposition 13.4.8,

### *13.4.9. Proposition. (Restriction of Representations)*

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a representation. Then there exist representations $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$, with commuting ranges, such that $\pi = \pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$. If $\pi$ is faithful, so are $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$. If $\pi$ is non-degenerate, then so are $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$.*

*Proof.* When $\pi$ is non-degenerate, use Exercises 13.4.17 and 13.4.18. For the degenerate case, apply the non-degenerate case to $\pi : \mathcal{A} \to \overline{\pi(\mathcal{A})\mathcal{H}}$. $\square$

When one goes the other way, from two representations $\pi_{\mathcal{A}}, \pi_{\mathcal{B}}$ with commuting ranges to $\pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$, it is possible for the product of faithful representations to not be faithful (Exercise 13.1.7). We use the notation $\pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$ to emphasize that the range is not necessarily inside a tensor product, as opposed to $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$.

Now we can define the main tensor norms.

### 13.4.10. Definition.

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras. If $\alpha$ is a C*-norm on $\mathcal{A} \otimes \mathcal{B}$, the **tensor product** $\mathcal{A} \otimes_\alpha \mathcal{B}$ is the C*-algebra obtained as the completion of $\mathcal{A} \otimes \mathcal{B}$ with respect to $\alpha$. We will mostly use the notation $\| \cdot \|_\alpha$ for such a norm.*

*The **maximal C*-norm** on $\mathcal{A} \otimes \mathcal{B}$ is given by*

$$\|x\|_{\max} = \sup\{\|\pi(x)\| : \ \pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H}) \ \text{ a representation}\}.$$

*The **minimal C*-norm** on $\mathcal{A} \otimes \mathcal{B}$ is given by*

$$\|x\|_{\min} = \|(\pi \otimes \rho)(x)\|_{\mathcal{B}(\mathcal{H} \otimes \mathcal{K})},$$

*where $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\rho : \mathcal{B} \to \mathcal{B}(\mathcal{K})$ are faithful representations.*

There are many things to unpack out of this definition. It is not entirely clear that the maximal and minimal norms are norms. It is not clear that $\|x\|_{\max}$ is not infinity, and it is not clear that the number $\|x\|_{\min}$ does not depend on the representations. All this will be addressed soon. One thing that is apparent from the definitions is that $\|x\|_{\min} \leq \|x\|_{\max}$ for all $x \in \mathcal{A} \otimes \mathcal{B}$.

*13.4.11. Remark*. There are varied schools of thought regarding notation and nomenclature for tensor products of C*-algebras. Although "max norm" and "min norm" are widely understood terms, the maximal and minimal norms are also called the **projective C*-norm** and **injective C*-norm** respectively; this latter terminology goes back to Grothendieck [Gro55]. A common name for the minimal norm is also the **spatial norm**. There is also the notation for the tensor product; whereas the notation $\mathcal{A} \otimes_{\max} \mathcal{B}$ is standard, some authors (notably Brown–Ozawa [BO08]) prefer to use $\mathcal{A} \odot \mathcal{B}$ for the algebraic tensor product and reserve $\mathcal{A} \otimes \mathcal{B}$ for the minimal C*-tensor product.

*13.4.12. Remark*. To mention an easy case that appears often, given a C*-algebra $\mathcal{A}$ and $n \in \mathbb{N}$ there is a single C*-norm on $M_n(\mathbb{C}) \otimes \mathcal{A}$. This follows from the fact that there is a canonical $*$-isomorphism $\gamma : M_n(\mathbb{C}) \otimes \mathcal{A} \to M_n(\mathcal{A})$, given by $\gamma(E_{kj} \otimes a)$ the matrix that contains $a$ in the $k, j$ entry and zeros elsewhere. We know from Section 11.7 that $M_n(\mathcal{A})$ is a C*-algebra, and so $\|x\|_m = \|\gamma(x)\|$ defines a C*-norm on $M_n(\mathbb{C}) \otimes \mathcal{A}$. If we embed $M_n(\mathbb{C}) \otimes \mathcal{A}$ into some C*-algebra (for instance $\mathcal{B}(\mathbb{C}^n \otimes \mathcal{K})$ for appropriate $\mathcal{K}$), $\gamma^{-1} : M_n(\mathcal{A}) \to M_n(\mathbb{C}) \otimes \mathcal{A}$ is a $*$-monomorphism and hence its image is closed by Proposition 11.4.9. So $M_n(\mathbb{C}) \otimes \mathcal{A}$ is a C*-algebra, isomorphic to $M_n(\mathcal{A})$.

For most purposes one can work with unital C*-algebras when discussing tensor products, for every C*-norm on $\mathcal{A} \otimes \mathcal{B}$ is the restriction of a C*-norm on $\tilde{\mathcal{A}} \otimes \tilde{\mathcal{B}}$. Indeed,

***13.4.13. Proposition.***

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras with $\mathcal{A}$ non-unital, and $\alpha$ a C*-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\alpha$ extends uniquely to a C*-norm in $\tilde{\mathcal{A}} \otimes \mathcal{B}$. In consequence, when both $\mathcal{A}$ and $\mathcal{B}$ are non-unital, a norm $\alpha$ on $\mathcal{A} \otimes \mathcal{B}$ extends uniquely to $\tilde{\mathcal{A}} \otimes \tilde{\mathcal{B}}$.*

*Proof.* Let $\pi : \mathcal{A} \otimes_\alpha \mathcal{B} \to \mathcal{B}(\mathcal{H})$ be a faithful representation (such $\pi$ always exists by Corollary 11.6.6). Then $\|x\|_\alpha = \|\pi(x)\|$ for all $x \in \mathcal{A} \otimes \mathcal{B}$, by Proposition 11.4.9. We may assume that $\pi$ is non-degenerate, via replacing $\mathcal{H}$ with $\overline{\pi(\mathcal{A} \otimes \mathcal{B})\mathcal{H}}$. By Proposition 13.4.9 we have faithful representations $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ with commuting ranges and $\pi = \pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$, and by Exercise 11.6.10 $\pi_{\mathcal{A}}$ can extended to a faithful representation $\tilde{\pi}_{\mathcal{A}} : \tilde{\mathcal{A}} \to \mathcal{B}(\mathcal{H})$. Let $\tilde{\pi} : \tilde{\mathcal{A}} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ be given by $\tilde{\pi}((a, \lambda) \otimes b) = \tilde{\pi}_{\mathcal{A}}(a, \lambda)\pi_{\mathcal{B}}(b)$

and extended by linearity. This is a representation (the multiplicativity at the level of elementary tensors guarantees that it is multiplicative on sums of tensors); if we show it is faithful, we can use it to define a norm $\tilde{\alpha}$ on $\tilde{\mathcal{A}} \otimes \mathcal{B}$ by

$$\|x\|_{\tilde{\alpha}} = \|\tilde{\pi}(x)\|, \qquad x \in \tilde{\mathcal{A}} \otimes \mathcal{B};$$

and when $x \in \mathcal{A} \otimes \mathcal{B}$ we have

$$\|x\|_{\tilde{\alpha}} = \|\tilde{\pi}(x)\| = \|\pi(x)\| = \|x\|_{\alpha}.$$

So let us prove that $\tilde{\pi}$ is faithful. Suppose that $\tilde{\pi}(x) = 0$ for some $x \in \tilde{\mathcal{A}} \otimes B$. We can write

$$x = \sum_{k=1}^{m} \tilde{a}_k \otimes b_k,$$

with $b_1, \ldots, b_m$ linearly independent. For approximate units $\{e_s\}$ of $\mathcal{A}$ and $\{f_t\}$ of $\mathcal{B}$, since $\tilde{\pi}$ is a homomorphism we have

$$\pi((e_s \otimes f_t)x) = \tilde{\pi}((e_s \otimes f_t)x) = \tilde{\pi}(e_s \otimes f_t)\tilde{\pi}(x) = 0$$

so, as $\pi$ is faithful,

$$0 = \sum_{k=1}^{m} e_s\tilde{a}_k \otimes f_t b_k.$$

By Exercise 11.4.8 $f_t b_1, \ldots, f_t b_m$ are linearly independent for large enough $t$, and hence $e_s\tilde{a}_k = 0$ for all $s$ and $k$. Writing $\tilde{a}_k = (a_k, \lambda_k)$, we have $0 = e_s\tilde{a}_k = (e_s a_k, 0)$, so $e_s a_k = 0$ for all $k$, and taking limit we get $a_k = 0$. Then $\tilde{a}_k = \lambda_k I_{\tilde{\mathcal{A}}}$ for $k = 1, \ldots, m$. With this information available, now

$$0 = \tilde{\pi}(x) = \tilde{\pi}\Big(\sum_{k=1}^{m} \lambda_k I_{\tilde{\mathcal{A}}} \otimes b_k\Big) = \sum_{k=1}^{n} \lambda_k \pi_{\mathcal{B}}(b_k) = \pi_{\mathcal{B}}\Big(\sum_{k=1}^{n} \lambda_k b_k\Big).$$

From the fact that $\pi_{\mathcal{B}}$ is faithful and $b_1, \ldots, b_m$ linearly independent we have that $\lambda_1 = \cdots = \lambda_m = 0$, and hence $x = 0$. This shows that $\tilde{\pi}$ is faithful.

We will defer the proof of uniqueness to Corollary 13.4.27, since we will need the minimality of the min norm.

When both $\mathcal{A}$ and $\mathcal{B}$ are non-unital, we perform the above twice. ☐

A consequence of being able to work on the unitization is the following. We will eventually learn that the inequality below is an equality, but that will require more machinery than we have available right now.

***13.4.14. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras, and $\alpha$ a C*-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\|a \otimes b\|_{\alpha} \leq \|a\|\,\|b\|$ for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$.*

*Proof.* By working on the unitizations if needed, we have

$$\|a \otimes b\|_\alpha = \|(a \otimes I_{\mathcal{B}})(I_{\mathcal{A}} \otimes b)\|_\alpha\| \leq \|a \otimes I_{\mathcal{B}}\|_\alpha \|I_{\mathcal{A}} \otimes b\|_\alpha.$$

We will show that $\|a \otimes I_{\mathcal{B}}\|_\alpha = \|a\|$, and with an analog argument for $b$ we conclude. To see that $\|a \otimes I_{\mathcal{B}}\|_\alpha = \|a\|$ we simply notice that $\gamma : a \longmapsto \|a \otimes I_{\mathcal{B}}\|$ is a C*-norm on $\mathcal{A}$. Indeed, we have $\gamma(a) \geq 0$ for all $a$. If $\gamma(a) = 0$, then $a \otimes I_{\mathcal{A}} = 0$ and hence $a = 0$ by Remark 13.1.5. The homogeneity and triangle inequality are clear, and

$$\gamma(a^* a) = \|a^* a \otimes I_{\mathcal{B}}\|_\alpha = \|(a \otimes I_{\mathcal{B}})^*(a \otimes I_{\mathcal{B}})\|_\alpha = \|a \otimes I_{\mathcal{B}}\|_\alpha^2 = \gamma(a)^2.$$

By (11.4), $\|a\| = \gamma(a) = \|a \otimes I_{\mathcal{A}}\|$. □

**13.4.5. The Max Norm.** The first observation we make is that $\|x\|_{\max}$ does exist for all $x \in \mathcal{A} \otimes \mathcal{B}$. Indeed, if $x = \sum_{k=1}^{m} a_k \otimes b_k$ and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ is a representation, using Proposition 13.4.9

$$\|\pi(x)\| \leq \sum_{k=1}^{m} \|\pi(a_k \otimes b_j)\| = \sum_{k=1}^{m} \|\pi_{\mathcal{A}}(a_k)\pi_{\mathcal{B}}(b_k)\| \leq \sum_{k=1}^{m} \|a_k\| \, \|b_k\| < \infty.$$

The triangle inequality follows from $\|\pi(x + y)\| \leq \|\pi(x)\| + \|\pi(y)\|$ and the subadditivity of the supremum. And the C*-identity works since

$$\|x\|_{\max}^2 = \sup\{\|\pi(x)\|^2 : \ \pi\} = \sup\{\pi(x^* x) : \ \pi\} = \|x^* x\|_{\max}.$$

So $\| \cdot \|_{\max}$ is a seminorm, and it remains to show that $\|x\|_{\max} = 0$ implies that $x = 0$. If $x \in \mathcal{A} \otimes \mathcal{B}$ is nonzero fix faithful representations $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\rho : \mathcal{B} \to \mathcal{B}(\mathcal{K})$. Then by Proposition 13.4.8 we have a faithful representation $\pi \otimes \rho : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$. So $(\pi \otimes \rho)(x) \neq 0$; hence $\|x\|_{\max} \geq \|(\pi \otimes \rho)(x)\| > 0$.

The "maximal" name is deserved. Almost directly from the definition, we get

***13.4.15. Proposition. (Universality of the max tensor product)***

*Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{C}$ a $*$-homomorphism. Then there exists a unique $*$-homomorphism $\hat{\pi} : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{C}$ that extends $\pi$. In particular, if $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{C}$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{C}$ are $*$-homomorphisms with commuting ranges, there exists a unique $*$-homomorphism extension $\pi_{\mathcal{A}} \times \pi_{\mathcal{B}} : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{C}$.*

*Proof.* We can represent $\mathcal{C} \subset \mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. Then $\pi$ becomes a representation of $\mathcal{A} \otimes \mathcal{B}$, and by definition we have $\|\pi(x)\| \leq \|x\|_{\max}$, so by density $\pi$ can by extended uniquely to $\mathcal{A} \otimes_{\max} \mathcal{B}$ via Proposition 6.1.9.

When $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ are $*$-homomorphisms with commuting ranges we construct $\pi_{\mathcal{A}} \times \pi_{\mathcal{A}} : \mathcal{A} \otimes \mathcal{B} \to \mathcal{C}$ via Corollary 13.1.9 and extend as above. $\square$

***13.4.16. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\alpha$ a $C^*$-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\|x\|_\alpha \leq \|x\|_{\max}$ for all $x \in \mathcal{A} \otimes_\alpha \mathcal{B}$.*

*Proof.* Let $\gamma : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ be the identity map. If we consider $\|\cdot\|_{\max}$ on the domain and $\|\cdot\|_\alpha$ on the codomain, by Proposition 13.4.15 we get a $*$-homomorphism $\hat{\gamma} : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{A} \otimes_\alpha \mathcal{B}$. As it has dense range, $\hat{\gamma}$ is surjective by Proposition 11.4.9. Then

$$\|x\|_\alpha = \|\hat{\gamma}(x)\|_\alpha \leq \|x\|_{\max}, \qquad x \in \mathcal{A} \otimes_\alpha \mathcal{B}$$

since $*$-homomorphisms are contractive. $\square$

We finish our very brief discussion of the max norm with the fact that products of completely positive maps work.

***13.4.17. Lemma.***

*Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ be $C^*$-algebras and $\psi_1 : \mathcal{A} \to \mathcal{C}$, $\psi_2 : \mathcal{B} \to \mathcal{D}$ completely positive maps. Then $\psi_1 \otimes \psi_2$ extends to a completely positive map $\psi_1 \otimes_{\max} \psi_2 : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{C} \otimes_{\max} \mathcal{D}$, with $\|\psi_1 \otimes_{\max} \psi_2\|_{\mathrm{cb}} = \|\psi_1\| \, \|\psi_2\|$.*

*Proof.* Since we can write $\psi_1 \otimes \psi_2 = (\psi_1 \otimes \mathrm{id}_{\mathcal{B}}) \circ (\mathrm{id}_{\mathcal{A}} \otimes \psi_2)$, we may assume without loss of generality that $\mathcal{D} = \mathcal{B}$ and $\psi_2 = \mathrm{id}_{\mathcal{B}}$. By fixing a faithful representation $\mathcal{C} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ we may assume that $\mathcal{C}, \mathcal{B} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{B} \subset \mathcal{C}' \subset \psi_1(\mathcal{A})'$. We write the minimal Stinespring Dilation

$$\psi_1(A) = V_1^* \pi_1(A) V_1, \qquad A \in \mathcal{A},$$

where $\pi_1 : \mathcal{A} \to \mathcal{B}(\mathcal{H}_1)$ is a representation and $V_1 : \mathcal{H}_1 \to \mathcal{H}$ is a bounded operator. By Exercise 13.2.17 there exists a $*$-homomorphism $\rho : \psi_1(\mathcal{A})' \to \pi_1(\mathcal{A})'$ with

$$\psi_1(A)T = V_1^* \pi_1(A) \rho(T) V_1, \qquad A \in \mathcal{A},\ T \in \psi_1(\mathcal{A})'.$$

Using Proposition 13.4.15 we get a representation $\pi_1 \otimes_{\max} \rho : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{B}(\mathcal{H})$ that satisfies $(\pi_1 \otimes \rho)(A \otimes B) = \pi_1(A)\rho(B)$. It follows that for $X = \sum_j A_j \otimes B_j \in \mathcal{A} \otimes \mathcal{B}$,

$$\begin{aligned} \|(\psi_1 \otimes \mathrm{id}_{\mathcal{B}})(X)\|_{\max} &= \Big\| \sum_j \psi_1(A_j) B_j \Big\|_{\max} = \Big\| \sum_j V^* \pi_1(A_j) \rho(B_j) V \Big\|_{\max} \\ &= \|V^*(\pi_1 \otimes_{\max} \rho)(X) V\|_{\max} \end{aligned}$$

$$\leq \|V\|^2 \,\|\pi_1 \otimes_{\max} \rho\| \,\|X\|_{\max}$$
$$= \|\psi_1\| \,\|X\|_{\max}.$$

This shows that $\psi_1 \otimes \mathrm{id}_{\mathcal{B}}$ is max-max bounded, and hence so is $\psi_1 \otimes \psi_2$. We also have

$$\|\psi_1 \otimes \psi_2\|_{\mathrm{cb}} = \|(V_1 \otimes V_2)^*(\pi_1 \otimes \pi_2)(V_1 \otimes V_2)\|$$
$$\leq \|V_1 \otimes V_2\|^2 = \|V_1\|^2\|V_2\|^2 = \|\psi_1\| \,\|\psi_2\|.$$

Conversely, if $\varepsilon > 0$ and $A \in \mathcal{A}, B \in \mathcal{B}$ with $\|A\| = \|B\| = 1$ and $\|\psi_1(A)\| \geq \|\psi_1\| - \varepsilon$, $\|\psi_2(B)\| \geq \|\psi_2\| - \varepsilon$, then

$$\|(\psi_1 \otimes \psi_2)(A \otimes B)\| = \|\psi_1(A)\| \,\|\psi_2(B)\| \geq \|\psi_1\| \,\|\psi_2\| - \varepsilon(\|\psi_1\| + \|\psi_2\|) - \varepsilon^2,$$

which establishes the reverse inequality. □

***13.4.18. Lemma.***

*Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be $C^*$-algebras and $\psi_1 : \mathcal{A} \to \mathcal{C}$, $\psi_2 : \mathcal{B} \to \mathcal{C}$ be completely positive maps with commuting ranges. Then $\psi_1 \times \psi_2$ extends to a completely positive map $\psi_1 \times_{\max} \psi_2 : \mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{C}$.*

*Proof.* Let $\mathcal{A}_1 = C^*(\psi_1(\mathcal{A})) \subset \mathcal{C}$ and $\mathcal{A}_2 = C^*(\psi_2(\mathcal{B})) \subset \mathcal{C}$. By hypothesis these subalgebras commute with each other. By Lemma 13.4.17 the map $\psi_1 \otimes \psi_2 : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}_1 \otimes \mathcal{A}_2$ is max-bounded. Let $\pi : \mathcal{C} \to \mathcal{B}(\mathcal{H})$ be a faithful representation, and $\rho : \mathcal{A}_1 \otimes \mathcal{A}_2 \to \mathcal{C}$ the $*$-homomorphism induced by $\rho(a \otimes b) = ab$ (here is where we use that $\mathcal{A}_1$ and $\mathcal{A}_2$ commute, to apply Corollary 13.1.9). As $\pi \circ \rho$ is a representation of $\mathcal{A} \otimes \mathcal{B}$ and $\pi$ is isometric, we have that $\|\rho(x)\| = \|\pi(\rho(x))\| \leq \|x\|_{\max}$ for all $x \in \mathcal{A} \otimes \mathcal{B}$ by definition of the max norm. Now

$$\|(\psi_1 \times \psi_2)(x)\| = \|(\rho \circ (\psi_1 \otimes_{\max} \psi_2))(x)\| \leq \|(\psi_1 \otimes_{\max} \psi_2)(x)\|_{\max} \leq \|x\|_{\max}.$$

Being a composition of completely positive maps, $\psi_1 \times \psi_2$ is completely positive. □

Complete positivity is crucial in Lemma 13.4.18. Indeed, Let $\mathcal{H} = \ell^2(\mathbb{N})$ and $\mathcal{A} = \mathcal{B} = \mathcal{B}(\mathcal{H})$, $\mathcal{C} = \mathcal{B}(\mathcal{H}) \otimes_{\max} \mathcal{B}(\mathcal{H})$. Let $\psi_1 : \mathcal{A} \to \mathcal{C}$ and $\psi_2 : \mathcal{B} \to \mathcal{C}$ be given by $\psi_1(S) = S^\top \otimes I_{\mathcal{H}}$ and $\psi_2(T) = I_{\mathcal{H}} \otimes T$. Let $\{E_{kj}\}$ be matrix units in $\mathcal{B}(\mathcal{H})$. As in Example 13.2.21, let

$$A = \sum_{k,j=1}^{n} E_{jk} \otimes E_{kj} \in \mathcal{A} \otimes \mathcal{B}, \qquad B = \sum_{k,j=1}^{n} E_{kj} \otimes E_{kj} \in \mathcal{A} \otimes \mathcal{B}.$$

Then

$$(\psi_1 \times \psi_2)(A) = \sum_{k,j=1}^{n} \psi_1(E_{jk})\psi_2(E_{kj}) = \sum_{k,j=1}^{n} E_{kj} \otimes E_{kj} = B.$$

From the calculation in the example we see that if $P_n = \sum_{k=1}^{n} E_{kk}$,

$$\|A^*A\|_{\max} = \|P_n \otimes P_n\|_{\max} = \|P_n\|^2 = 1.$$

We also know that $\|B\| = n$. So $\psi_1 \times \psi_2$ is not max-bounded and hence cannot be extended to $\mathcal{A} \otimes_{\max} \mathcal{B}$.

**13.4.6. The Min Norm.** In the definition of the min norm, because $\pi$ and $\rho$ are faithful, so is $\pi \otimes \rho$ by Proposition 13.4.8, and this shows that min is a norm. What is not clear is that the definition does not depend on the actual faithful representations used. The approach in [Tak02, Theorem IV.4.9] is to show that

$$\|x\|_{\min} = \sup\left\{(\varphi \otimes \psi)(y^*x^*xy)^{1/2} : \ \varphi \in S(\mathcal{A}),\ \psi \in S(\mathcal{B}),\ (\varphi \otimes \psi)(y^*y) = 1\right\},$$

via Propositions 11.6.3 and 11.6.8 (and a lot of work after!). The proof here will in the end use the above equality, but the path we will follow is the one suggested in [BO08].

***13.4.19. Proposition.***

*The min norm does not depend on the choice of faithful representations used.*

*Proof.* It is enough to show that we can replace one representation, as then we can do the other. So assume we have $\|\cdot\|_{\min}$ defined in terms of faithful representations $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\rho : \mathcal{B} \to \mathcal{B}(\mathcal{K})$, and assume that $\rho' : \mathcal{B} \to \mathcal{B}(\mathcal{K}')$ is another faithful representation. We denote

$$\|x\|_{\min'} = \|(\pi \otimes \rho')(x)\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K}')}.$$

By Exercise 10.5.12 we have an increasing net $\{P_j\} \subset \mathcal{B}(\mathcal{H})$ of finite-rank projections with $P_j\xi \to \xi$ for all $\xi \in \mathcal{H}$. Then for any $T \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$, using Exercise 13.4.13 and Exercise 10.5.13,

$$\|T\| = \sup\{\|(P_j \otimes I_{\mathcal{K}})T(P_j \otimes I_{\mathcal{K}})\| : \ j\}.$$

This allows us to write (with the second to last equality to be justified)

$$\begin{aligned}\left\|\sum_{r=1}^{m} a_r \otimes b_r\right\|_{\min} &= \left\|\sum_{r=1}^{m} \pi(a_r) \otimes \rho(b_r)\right\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K})} \\ &= \sup_j \left\|\sum_{r=1}^{m} P_j\pi(a_r)P_j \otimes \rho(b_r)\right\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K})}\end{aligned}$$

$$= \sup_j \left\| \sum_{r=1}^m P_j\pi(a_r)P_j \otimes \rho'(b_r) \right\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K})}$$

$$= \left\| \sum_{r=1}^m a_r \otimes b_r \right\|_{\min'}.$$

The reason we can switch from $\rho$ to $\rho'$ is that $P_j\mathcal{B}(\mathcal{H})P_j$ can be naturally identified with $\mathcal{B}(P_j\mathcal{H}) \simeq M_n(\mathbb{C})$, where $n = \operatorname{Tr}(P_j) = \dim P_j\mathcal{H}$. If $\alpha$ is a $*$-isomorphism between $P_j\mathcal{B}(\mathcal{H})P_j$ and $M_n(\mathbb{C})$, isometric by Proposition 11.4.9, then $\alpha \otimes \rho^{-1}$ is a $*$-isomorphism between $P_j\mathcal{B}(\mathcal{H})P_j \otimes \rho(\mathcal{B})$ and $M_n(\mathbb{C}) \otimes \mathcal{B}$ (injective by Proposition 13.4.8, with dense range so surjective by Proposition 11.4.9); and $\alpha \otimes (\rho')^{-1}$ is another such $*$-isomorphism. As $M_n(\mathbb{C}) \otimes \mathcal{B}$ has a unique C$^*$-norm by Remark 13.4.12, we get

$$\left\| \sum_{r=1}^m P_j\pi(a_r)P_j \otimes \rho(b_r) \right\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K})} = \left\| \sum_{r=1}^m \alpha(P_j\pi(a_r)P_j) \otimes b_r \right\|_{M_n(\mathbb{C})\otimes\mathcal{B}}$$

$$= \left\| \sum_{r=1}^m P_j\pi(a_r)P_j \otimes \rho'(b_r) \right\|_{\mathcal{B}(\mathcal{H}\otimes\mathcal{K}')}. \qquad \square$$

Now we build towards Corollary 13.4.21.

***13.4.20. Lemma.***

*Let $\mathcal{A}$, $\mathcal{B}$ be unital abelian C$^*$-algebras, $\alpha$ a C$^*$-norm on $\mathcal{A} \otimes \mathcal{B}$, and $\rho \in \Sigma(\mathcal{A})$, $\gamma \in \Sigma(\mathcal{B})$. Then $\rho \otimes \gamma : \mathcal{A} \otimes \mathcal{B}$ extends to a state on $\mathcal{A} \otimes_\alpha \mathcal{B}$.*

*Proof.* Consider the set

$$\mathcal{U} = \{(\varphi, \psi) \in \Sigma(\mathcal{A}) \times \Sigma(\mathcal{B}) : \ |\varphi \otimes \psi(x)| \leq \|x\|_\alpha, \ x \in \mathcal{A} \otimes \mathcal{B}\}.$$

We want to show that $\mathcal{U} = \Sigma(\mathcal{A}) \times \Sigma(\mathcal{B})$. We consider the product weak$^*$-topology on $\Sigma(\mathcal{A}) \times \Sigma(\mathcal{B})$. As this is pointwise convergence and the elements of $\mathcal{A} \otimes \mathcal{B}$ are finite sums, $\mathcal{U}$ is closed. If $\mathcal{U} \subsetneq \Sigma(\mathcal{A}) \times \Sigma(\mathcal{B})$, as the complement is open we can find a basic open rectangle $U_\mathcal{A} \times U_\mathcal{B}$ in $\Sigma(\mathcal{A}) \times \Sigma(\mathcal{B})$, disjoint with $\mathcal{U}$. Via the Gelfand transform we see $\mathcal{A}$ as the continuous functions on $\Sigma(\mathcal{A})$; as $U_\mathcal{A}$ is open, we can choose $a \in \mathcal{A}$ with $0 \leq a \leq I_\mathcal{A}$ and its support as a function inside $U_\mathcal{A}$. Similarly we choose $b \in \mathcal{B}$ with $0 \leq b \in I_\mathcal{B}$ and $b$ supported inside $U_\mathcal{B}$. This means that $(\varphi \otimes \psi)(a \otimes b) = 0$ for all $(\varphi, \psi) \in \mathcal{U}$. Since $\mathcal{A} \otimes_\alpha \mathcal{B}$ is abelian, there exists a character $\eta \in \Sigma(\mathcal{A} \otimes_\alpha \mathcal{B})$ with $\eta(a \otimes b) > 0$ (via the Gelfand transform since $a \otimes b$ is a non-negative nonzero function on $\Sigma(\mathcal{A} \otimes_\alpha \mathcal{B})$.

By Corollary 13.2.40 the GNS representation $\pi_\eta : \mathcal{A} \otimes_\alpha \mathcal{B} \to \mathcal{B}(H_\eta)$ is irreducible. As $\mathcal{A} \otimes_\alpha \mathcal{B}$ is abelian, this means that $\dim \mathcal{H}_\eta = 1$. Hence the representation $\pi_{\eta,\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\eta)$ given by $\pi_{\eta,\mathcal{A}}(a) = \pi_\eta(a \otimes I_\mathcal{B})$ is irreducible. This is a minimal Stinespring representation for the state $\eta_\mathcal{A}(a) = \eta(a \otimes I_\mathcal{B})$, since

$$\eta_\mathcal{A}(a) = \eta(a \otimes I_\mathcal{B}) = V\pi_\eta(a \otimes I_\mathcal{B})V^* = V\pi_{\eta,\mathcal{A}}(a)V*\cdot$$

The one-dimensionality guarantees that the dilation is minimal. As the representation is also irreducible, $\eta|_\mathcal{A}$ is pure by Corollary 13.2.40; similarly, $\eta|_\mathcal{B}$ is pure. This gives us a contradiction, for

$$\big|\,\eta|_\mathcal{A} \otimes \eta|_\mathcal{B}(x)\,\big| = |\eta(x)| \leq \|x\|_\alpha, \qquad x \in \mathcal{A} \otimes \mathcal{B}$$

so $\eta|_\mathcal{A} \otimes \eta|_\mathcal{B} \in \mathcal{U}$, while $\eta|_\mathcal{A} \otimes \eta|_\mathcal{B}(a \otimes b) = \eta(a \otimes b) > 0$.

Now $\rho \otimes \gamma$ extends to a state on $\mathcal{A} \otimes_\alpha \mathcal{B}$ by Exercise 11.5.5. □

***13.4.21. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras, $\alpha$ a $C^*$-norm on $\mathcal{A} \otimes \mathcal{B}$, and $\varphi \in S(\mathcal{A})$, $\psi \in S(\mathcal{B})$. Then $\varphi \otimes \psi$ is bounded for $\alpha$.*

*Proof.* By Proposition 13.4.13 and Proposition 11.5.6, we may assume without loss of generality that $\mathcal{A}$ and $\mathcal{B}$ are unital. Assume first that $\varphi$ and $\psi$ are pure. By Proposition 13.2.48 there exist nets $\{e_j\} \subset \mathcal{A}$ and $\{f_k\} \subset \mathcal{B}$ with $\varphi(e_j) = 1$ for all $j$, $\psi(f_k) = 1$ for all $k$, $\lim_j \varphi(a)e_j^2 - e_j a e_j = 0$ for all $a \in \mathcal{A}$, and $\lim_k \psi(b)f_k^2 - f_k b f_k = 0$ for all $b \in \mathcal{B}$.

Given $x = \sum_{r=1}^{m} a_r \otimes b_r \in \mathcal{A} \otimes \mathcal{B}$ and using Corollary 13.4.14

$$\begin{aligned}
&\|(e_j \otimes f_k)x(e_j \otimes f_k) - (\varphi \otimes \psi)(x)(e_j \otimes f_k)^2\|_\alpha \\
&\qquad \leq \sum_r \|e_j a_r e_j \otimes f_k b_r f_k - \varphi(a_r)e_j^2 \otimes \psi(b_r)f_k^2\|_\alpha \\
&\qquad \leq \sum_r \|(e_j a_r e_j - \varphi(a_r)e_j^2) \otimes f_k b_r f_k\|_\alpha \\
&\qquad\qquad + \|\varphi(a_r)e_j^2 \otimes (f_k b_r f_k - \psi(b_r)f_k^2)\|_\alpha \\
&\qquad \leq \sum_r \|e_j a_r e_j - \varphi(a_r)e_j^2|\,\|f_k b_r f_k\| + \|\varphi(a_r)e_j^2\|\,\|f_k b_r f_k - \psi(b_r)f_k^2\| \\
&\qquad \leq \sum_r c\|e_j a_r e_j - \varphi(a_r)e_j^2\| + \|\varphi\|\,c\,\|f_k b_r f_k - \psi(b_r)f_k^2\| \\
&\qquad \xrightarrow[j,k]{} 0.
\end{aligned}$$

We have $\|e_j\| = 1$, for $1 = \varphi(e_j) \leq \|e_j\|$ and $e_j \leq I_\mathcal{A}$; and similarly $\|f_k\| = 1$ for all $k$. We claim that $\|e_j \otimes f_k\|_\alpha = 1$. For this, we already know that $\|e_j \otimes f_k\|_\alpha \leq \|e_j\|\,\|f_k\| = 1$ by Corollary 13.4.14. The $*$-subalgebra $C^*(e_j, I_\mathcal{A}) \otimes$

$C^*(f_k, I_\mathcal{B})$ is abelian and it contains the unit $I_\mathcal{A} \otimes I_\mathcal{B}$. By Theorem 11.2.3 there exist characters $\rho : C^*(e_j, I_\mathcal{A}) \to \mathbb{C}$ and $\gamma : C^*(f_k, I_\mathcal{B}) \to \mathbb{C}$ such that $\rho(e_j) = \gamma(f_k) = 1$. By Lemma 13.4.20, $\rho \otimes \gamma : C^*(e_j, I_\mathcal{A}) \otimes C^*(f_k, I_\mathcal{B}) \to \mathbb{C}$ is a bounded character. This can be extended by Hahn–Banach (Corollary 5.7.6) to a linear functional $\tilde{\eta}$ on $\mathcal{A} \otimes_\alpha \mathcal{B}$; as this is unital and has norm equal to 1, it is positive by Proposition 11.5.4, so a state. We have $\tilde{\eta}(e_j \otimes f_k) = \rho(e_j)\psi(f_k) = 1$, and therefore $\|e_j \otimes f_k\|_\alpha = 1$ (as we already knew that it was a contraction). As $e_j \otimes f_k$ is positive, we also get $\|(e_j \otimes f_k)^2\|_\alpha = \|e_j \otimes f_k\|_\alpha^2 = 1$. Then (note that the first equality is true for all $j, k$ even without limit)

$$\begin{aligned} |(\varphi \otimes \psi)(x)| &= \lim_{j,k} \|(\varphi \otimes \psi)(x)(e_j \otimes f_k)^2\|_\alpha \\ &= \lim_{j,k} \|(e_j \otimes f_k)x(e_j \otimes f_k)\|_\alpha \leq \|x\|_\alpha. \end{aligned}$$

When $\varphi$ is not pure, combining Proposition 13.2.41 with Proposition 11.5.15 and with Theorem 7.5.11 we have that $\varphi = \lim_h \varphi_h$ pointwise (that is, in the weak*-topology) with each $\varphi_h$ a convex combination of pure states. For each $\varphi_h = \sum_s t_s \varphi_{hs}$ with $\varphi_{hs}$ pure, by the estimate above

$$|(\varphi_h \otimes \psi)(x)| \leq \sum_s t_s |(\varphi_{hs} \otimes \psi)(x)| \leq \sum_s t_s \|x\|_\alpha = \|x\|_\alpha,$$

and then

$$|(\varphi \otimes \psi)(x)| = \lim_h |(\varphi_h \otimes \psi)(x)| \leq \|x\|_\alpha.$$

We can now get the general statement by proceeding similarly with $\psi$. □

The remarkable thing about Corollary 13.4.21 lies in the fact that we get that product states are bounded for any norm. This will allow us to show that the min norm deserves its name.

**13.4.22. Theorem. (Takesaki)**

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\alpha$ a $C^*$-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\|x\|_{\min} \leq \|x\|_\alpha$ for all $x \in \mathcal{A} \otimes \mathcal{B}$.*

*Proof.* By Proposition 13.4.13 we may assume without loss of generality that $\mathcal{A}$ and $\mathcal{B}$ are unital. Also, as it is enough to prove the inequality in $C^*(x)$ for $x = \sum_j a_j \otimes b_j$, we may replace $\mathcal{A}$ and $\mathcal{B}$ by $C^*(a_1, \dots, a_n)$ and $C^*(b_1, \dots, b_n)$ respectively; so we may assume that $\mathcal{A}$ and $\mathcal{B}$ are separable. Then Proposition 11.5.12 gives us faithful states $\varphi \in S(\mathcal{A})$ and $\psi \in S(\mathcal{B})$. From this we get faithful GNS representations $\pi_\varphi$ and $\pi_\psi$ (Exercise 11.6.6), and so from Proposition 13.4.19

$$\|x\|_{\min} = \|(\pi_\varphi \otimes \pi_\psi)(x)\|_{\mathcal{B}(\mathcal{H}_\varphi \bar{\otimes} \mathcal{H}_\psi)}.$$

This equality in particular implies that $\pi_\varphi \otimes \pi_\psi$ extends to a representation $\mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{B}(\mathcal{H}_\varphi \otimes \mathcal{H}_\psi)$, which is cyclic (Exercise 13.4.20). We also have that $\varphi \otimes \psi$ is a bounded functional on $\mathcal{A} \otimes_\alpha \mathcal{B}$ by Corollary 13.4.21. By doing GNS we have a representation $\pi_{\varphi\otimes_\alpha\psi} : \mathcal{A} \otimes_\alpha \mathcal{B} \to \mathcal{B}(\mathcal{H}_{\varphi\otimes\psi})$. As

$$\begin{aligned}\langle(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(a \otimes b)(\xi_{\mathcal{A}} \otimes \xi_{\mathcal{B}}), \xi_{\mathcal{A}} \otimes \xi_{\mathcal{B}}\rangle &= \langle\pi_{\mathcal{A}}(a)\xi_{\mathcal{A}}, \xi_{\mathcal{A}}\rangle \langle\pi_{\mathcal{B}}(b)\xi_{\mathcal{B}}, \xi_{\mathcal{B}}\rangle \\ &= \varphi(a)\psi(b) = (\varphi \otimes \psi)(a \otimes b),\end{aligned}$$

extending this equality by linearity and continuity we get that

$$(\varphi \otimes \psi)(x) = \langle(\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}})(x)(\xi_{\mathcal{A}} \otimes \xi_{\mathcal{B}}), \xi_{\mathcal{A}} \otimes \xi_{\mathcal{B}}\rangle$$

for all $x \in \mathcal{A} \otimes \mathcal{B}$. By Proposition 11.6.3, the representations $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$ and $\pi_{\varphi\otimes\psi}$ are unitarily equivalent. It follows that for any $x \in \mathcal{A} \otimes \mathcal{B}$ we have (by Proposition 11.1.11)

$$\|x\|_{\min} = \|\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}(x)\| = \|\pi_{\varphi\otimes\psi}(x)\| \le \|x\|_\alpha. \qquad \square$$

As an immediate consequence of Takeksaki's Theorem,

***13.4.23. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\alpha$ a $C^*$-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\alpha$ is a cross norm.*

*Proof.* Since the min norm is a cross norm, by Theorem 13.4.22 and Corollary 13.4.14

$$\|a\|\,\|b\| = \|a \otimes b\|_{\min} \le \alpha(a \otimes b) \le \|a\|\,\|b\|. \qquad \square$$

***13.4.24. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be $C^*$-algebras and $\pi : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a representation. If $\pi$ is faithful when restricted to $\mathcal{A} \otimes \mathcal{B}$, then $\pi$ is faithful.*

*Proof.* Because $\pi$ is faithful on $\mathcal{A} \otimes \mathcal{B}$, we can define a norm $\alpha(x) = \|\pi(x)\|$ on $\mathcal{A} \otimes \mathcal{B}$. Suppose that $x \in \mathcal{A} \otimes_{\min} \mathcal{B}$ with $\pi(x) = 0$. Let $\{x_n\} \subset \mathcal{A} \otimes \mathcal{B}$ with $x_n \to x$ for some $x \in \mathcal{A} \otimes_{\min} \mathcal{B}$. By Theorem 13.4.22,

$$0 = \|\pi(x)\| = \lim_n \|\pi(x_n)\| \ge \lim_n \|x_n\|_{\min} = \|x\|_{\min}.$$

Hence $x = 0$ and $\pi$ is faithful. $\square$

***13.4.25. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be simple $C^*$-algebras. Then $\mathcal{A} \otimes_{\min} \mathcal{B}$ is simple.*

*Proof.* Let $\pi : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{B}(\mathcal{H})$ be an irreducible representation. By Proposition 13.4.9 there exist representations $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$ with commuting ranges and such that $\pi = \pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$ on $\mathcal{A} \otimes \mathcal{B}$. Then, as $\mathcal{A} \otimes \mathcal{B}$ is norm-dense in $\mathcal{A} \otimes_{\min} \mathcal{B}$, and using the irreducibility,

$$\begin{aligned}\mathcal{B}(\mathcal{H}) &= \pi(\mathcal{A} \otimes_{\min} \mathcal{B})'' = \pi(\mathcal{A} \otimes \mathcal{B})'' \\ &= \{\pi_{\mathcal{A}}(\mathcal{A}) \cup \pi_{\mathcal{B}}(\mathcal{B})\}'' = \{\pi_{\mathcal{A}}(\mathcal{A})'' \cup \pi_{\mathcal{B}}(\mathcal{B})''\}''.\end{aligned}$$

As every element of $\pi(\mathcal{A})''$ is a pointwise limit of elements in $\pi(\mathcal{A})$ by the Double Commutant Theorem (12.3.5), we still have that every element of $\pi_{\mathcal{A}}(\mathcal{A})''$ commutes with every element of $\pi_{\mathcal{B}}(\mathcal{B})''$. Thus an element in the centre of $\pi_{\mathcal{A}}(\mathcal{A})''$ is in the centre of $\mathcal{B}(\mathcal{H})$, and so it is a scalar; that is, $M_{\mathcal{A}} = \pi_{\mathcal{A}}(\mathcal{A})''$ and $M_{\mathcal{B}} = \pi_{\mathcal{B}}(\mathcal{B})''$ are factors with $M_{\mathcal{A}} \subset M'_{\mathcal{B}}$, $M_{\mathcal{B}} \subset M'_{\mathcal{A}}$. By Proposition 13.1.11 the map $\rho : TS \longmapsto T \otimes S$ induces an isomorphism between $\operatorname{alg}\{M_{\mathcal{A}} \cup M_{\mathcal{B}}\}$ and $M_{\mathcal{A}} \otimes M_{\mathcal{B}}$. In particular $\rho \circ \pi : \mathcal{A} \otimes \mathcal{B} \to \pi_{\mathcal{A}}(\mathcal{A}) \otimes \pi_{\mathcal{B}}(\mathcal{B})$ agrees with $\pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$; each of $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ is faithful by $\mathcal{A}$ and $\mathcal{B}$ being simple, so $\rho \circ \pi$ is faithful on $\mathcal{A} \otimes \mathcal{B}$ by Proposition 13.4.9. But then $\rho \circ \pi$ is faithful on $\mathcal{A} \otimes_{\min} \mathcal{B}$ by Corollary 13.4.24, and then $\pi$ is faithful. As $\pi$ was an arbitrary irreducible representation of $\mathcal{A} \otimes_{\min} \mathcal{B}$, by Corollary 13.2.43, $\mathcal{A} \otimes_{\min} \mathcal{B}$ is simple. □

For any norm $\alpha$ that does not agree with min on $\mathcal{A} \otimes \mathcal{B}$, the tensor product $\mathcal{A} \otimes_{\alpha} \mathcal{B}$ cannot be simple; for the "simple" reason that the identity map $\mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$ extends to a $*$-homomorphism $\mathcal{A} \otimes_{\alpha} \mathcal{B} \to \mathcal{A} \otimes_{\min} \mathcal{B}$, and this $*$-homomorphism is surjective as it has dense range, which makes $\mathcal{A} \otimes_{\min} \mathcal{B}$ a quotient of $\mathcal{A} \otimes_{\alpha} \mathcal{B}$.

***13.4.26. Lemma.***

*Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ be $C^*$-algebras and $\psi_1 : \mathcal{A} \to \mathcal{C}$, $\psi_2 : \mathcal{B} \to \mathcal{D}$ completely positive maps. Then $\psi_1 \otimes \psi_2$ extends to a completely positive map $\psi_1 \otimes_{\min} \psi_2 : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{C} \otimes_{\min} \mathcal{D}$, with $\|\psi_1 \otimes \psi_2\|_{\mathrm{cb}} = \|\psi_1\| \, \|\psi_2\|$. If $\psi_1$ and $\psi_2$ are faithful then $\psi_1 \otimes_{\min} \psi_2$ is faithful.*

*Proof.* Assuming $\mathcal{C} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{D} \subset \mathcal{B}(\mathcal{K})$, we write $\psi_1, \psi_2$ in terms of their Stinespring Dilations:

$$\psi_1(a) = V_1 \pi_1(a) V_1^*, \qquad \psi_2(b) = V_2 \pi_2(b) V_2^*, \qquad a \in \mathcal{A}, b \in \mathcal{B},$$

where $\pi_1 : \mathcal{A} \to \mathcal{B}(\mathcal{H}_1)$ and $\pi_2 : \mathcal{B} \to \mathcal{B}(\mathcal{H}_2)$ are representations, and $V_1 : \mathcal{H}_1 \to \mathcal{H}$ and $V_1 : \mathcal{H}_2 \to \mathcal{K}$ are bounded operators. As mentioned in Theorem 13.2.15 we may also assume without loss of generality that $\pi_1, \pi_2$ are faithful. As $\pi_1, \pi_2$ are faithful we have $\|(\pi_1 \otimes \pi_2)(x)\|_{\mathcal{B}(\mathcal{H}_1 \bar{\otimes} \mathcal{H}_2)} = \|x\|_{\min}$

by definition, and so $\pi_1 \otimes \pi_2$—being isometric—extends to a faithful representation of $\mathcal{A} \otimes_{\min} \mathcal{B}$ into $\mathcal{B}(\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2)$. For $x \in \mathcal{A} \otimes \mathcal{B}$,

$$\begin{aligned}\|(\psi_1 \otimes \psi_2)(x)\|_{\mathcal{B}(\mathcal{H} \overline{\otimes} \mathcal{K})} &= \|(V_1 \otimes V_2)(\pi_1 \otimes \pi_2)(x)(V_1 \otimes V_2)^*\| \\ &\leq \|V_1 \otimes V_2\|^2 \, \|(\pi_1 \otimes \pi_2)(x)\| = \|V_1 \otimes V_2\|^2 \, \|x\|_{\min} \\ &= \|\psi_1\| \, \|\psi_2\| \, \|x\|_{\min}.\end{aligned}$$

So $\psi_1 \otimes \psi_2$ extends uniquely to a bounded map $\mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{C} \otimes_{\min} \mathcal{D}$ (since we can define the min norm on $\mathcal{C} \otimes \mathcal{D}$ by using the identity representation). The equality of the norm follows from $\|\psi_1 \otimes \psi_2\| = \|\psi_1\| \, \|\psi_2\|$ (obtained from the above inequality and approximating the norm with elementary tensors as in the proof of Lemma 13.4.17); then $\|\psi_1 \otimes \psi_2\|_{\mathrm{cb}} = \|\psi_1 \otimes \psi_2\|$ by Proposition 13.2.24.

As for the complete positivity, we know that $\psi_1 \otimes \psi_2 = (V_1 \otimes V_2)(\pi_1 \otimes \pi_2)(V_1 \otimes V_2)^*$ is completely positive on $\mathcal{A} \otimes \mathcal{B}$. For $x \in M_n(\mathcal{A} \otimes_{\min} \mathcal{A})$ we can write it as a norm limit of elements $x_n \in M_n(\mathcal{A} \otimes \mathcal{B})$, and so $(\psi_1 \otimes_{\min} \psi_2)^{(n)}(x) \geq 0$ as it is a limit of positive elements.

Now the faithfulness. Consider first the case where $\mathcal{C} = \mathbb{C}$, $\mathcal{D} = \mathcal{B}$, $\psi_2 = \mathrm{id}_{\mathcal{B}}$, and $\{\varphi_j\} \subset S(\mathcal{A})$ is a faithful family of states (that is, $\varphi_j(x^*x) = 0$ for all $j$ implies $x = 0$). Let $x \in \mathcal{A} \otimes_{\min} \mathcal{B}$ such that $(\varphi_j \otimes \mathrm{id}_{\mathcal{B}})(x^*x) = 0$ for all $j$. Given a state $\eta \in S(\mathcal{B})$, we have

$$(\varphi_j \otimes \mathrm{id}_{\mathbb{C}})(\mathrm{id}_{\mathcal{A}} \otimes \eta)(x^*x) = (\mathrm{id}_{\mathcal{A}} \otimes \eta)(\varphi_j \otimes \mathrm{id}_{\mathcal{B}})(x^*x) = 0.$$

Since $\mathcal{A} \otimes \mathbb{C} \simeq \mathcal{A}$ canonically, we can identify $\varphi_j \otimes \mathrm{id}_{\mathbb{C}}$ with $\varphi_j$ and then the faithfulness of the family gives us that $(\mathrm{id}_{\mathcal{A}} \otimes \eta)(x^*x) = 0$. For a state $\varphi \in S(\mathcal{A})$,

$$(\varphi \otimes \eta)(x^*x) = (\varphi \otimes \mathrm{id}_{\mathbb{C}})(\mathrm{id}_{\mathcal{A}} \otimes \eta)(x^*x) = 0.$$

As $\varphi$ and $\eta$ were arbitrary, Exercise 13.4.22 gives us that $x = 0$. Next consider the case where $\mathcal{D} = \mathcal{B}$ and $\psi_2 = \mathrm{id}_{\mathcal{B}}$. As $\psi_1$ is faithful, the family $\{\varphi \circ \psi_1 : \varphi \in S(\mathcal{C})\}$ is faithful since states separate points by Corollary 11.5.8. If $(\psi_1 \otimes \mathrm{id}_{\mathcal{B}})(x^*x) = 0$, then for any $\varphi \in S(\mathcal{C})$

$$(\varphi \circ \psi_1 \otimes \mathrm{id}_{\mathcal{B}})(x^*x) = (\varphi \otimes \mathrm{id}_{\mathcal{B}})(\psi_1 \otimes \mathrm{id}_{\mathcal{B}})(x^*x) = 0.$$

The previous case then gives us $x = 0$, so $\psi_1 \otimes \mathrm{id}_{\mathcal{B}}$ is faithful. Now, in the general case,

$$\psi_1 \otimes \psi_2 = (\psi_1 \otimes \mathrm{id}_{\mathcal{D}})(\mathrm{id}_{\mathcal{A}} \otimes \psi_2)$$

is faithful as it is a product of faithful maps. □

Complete positivity is really necessary in Lemma 13.4.26. For instance, one can rewrite Example 13.2.21 to show that if $\mathcal{A} = \mathcal{K}(\mathcal{H})$ with $\mathcal{H}$ infinite-dimensional and $\psi : \mathcal{A} \to \mathcal{A}$ is the transpose map, then $\psi$ is a positive isometry, but $\psi \otimes \mathrm{id}_{\mathcal{A}} : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$ is unbounded.

It is also important to remember that $*$-homomorphisms on a C*-algebra are completely positive (Exercise 13.2.3; see also Exercise 13.2.4), so Lemma 13.4.26 works for them.

We can now prove the uniqueness in Proposition 13.4.13.

***13.4.27. Corollary.***

*Let $\mathcal{A}, \mathcal{B}$ be C*-algebras with $\mathcal{A}$ non-unital, and $\alpha$ a C*-norm on $\mathcal{A} \otimes \mathcal{B}$. Then $\alpha$ extends uniquely to a C*-norm in $\tilde{\mathcal{A}} \otimes \mathcal{B}$. In consequence, when both $\mathcal{A}$ and $\mathcal{B}$ are non-unital, a norm $\alpha$ on $\mathcal{A} \otimes \mathcal{B}$ extends uniquely to $\tilde{\mathcal{A}} \otimes \tilde{\mathcal{B}}$.*

*Proof.* The existence was shown in Proposition 13.4.13. It remains to show the uniqueness, and as in the proof of Proposition 13.4.13 it is enough to work with one unitization. So suppose that $\alpha, \beta$ are norms in $\tilde{\mathcal{A}} \otimes \mathcal{B}$ such that $\alpha = \beta$ on $\mathcal{A} \otimes \mathcal{B}$. Let $\gamma = \max\{\alpha, \beta\}$; then $\gamma$ is a C*-norm on $\mathcal{A} \otimes \mathcal{B}$. Consider the map $\rho : \mathcal{A} \otimes_\gamma \mathcal{B} \to \mathcal{A} \otimes_\alpha \mathcal{B}$ obtained by extending the identity map (this works, and it makes $\rho$ bounded, because $\|\rho(x)\|_\alpha = \|x\|_\alpha \leq \|x\|_\gamma$). So $\rho$ is a $*$-homomorphism with dense range, and hence surjective. Let $\psi \in S(\tilde{\mathcal{A}})$ be the unique state with $\ker \psi = \mathcal{A}$ (Exercise 11.5.12). We have that $\psi \otimes_{\min} \mathrm{id}_{\mathcal{B}}$ is bounded and completely positive by Lemma 13.4.26. In particular, for $x \in \tilde{\mathcal{A}} \otimes \mathcal{B}$,

$$\|(\psi \otimes \mathrm{id}_{\mathcal{B}})(x)\| \leq \|x\|_{\min} \leq \|x\|_\alpha \leq \|x\|_\gamma.$$

and $(\psi \otimes \mathrm{id}_{\mathcal{B}})(\rho(x)) = (\psi \otimes \mathrm{id}_{\mathcal{B}})(x)$. As $\rho$ is $\gamma$–$\alpha$ continuous and $\psi \otimes \mathrm{id}_{\mathcal{B}}$ is both $\gamma$-continuous and $\alpha$-continuous, the equality extends to $\tilde{\mathcal{A}} \otimes_\gamma \mathcal{B}$:

$$(\psi \otimes \mathrm{id}_{\mathcal{B}})(\rho(x)) = (\psi \otimes \mathrm{id}_{\mathcal{B}})(x), \qquad x \in \tilde{\mathcal{A}} \otimes_\gamma \mathcal{B}. \tag{13.24}$$

Now let $x \in \tilde{\mathcal{A}} \otimes_\gamma \mathcal{B}$ with $\rho(x) = 0$. By (13.24) we have $(\psi \otimes \mathrm{id}_{\mathcal{B}})(x) = 0$, and hence $x \in \mathcal{A} \otimes_\gamma \mathcal{B}$ by Exercise 13.4.21. By hypothesis $\alpha = \beta$ on $\mathcal{A} \otimes \mathcal{B}$, so $\gamma = \alpha = \beta$ on $\mathcal{A} \otimes \mathcal{B}$, and thus $\mathcal{A} \otimes_\gamma \mathcal{B} = \mathcal{A} \otimes_\alpha \mathcal{B}$. This means that $0 = \|\rho(x)\|_\alpha = \|x\|_\alpha$ and hence $x = 0$. Therefore $\rho$ is a $*$-isomorphism $\tilde{\mathcal{A}} \otimes_\gamma \mathcal{B} \to \tilde{\mathcal{A}} \otimes_\alpha \mathcal{B}$ which implies that $\gamma = \alpha = \beta$. ☐

**13.4.7. Max and Min are not Necessarily Equal.** Let $\mathcal{A} = \mathcal{B} = C^*_\lambda(\mathbb{F}_2)$. The original example of two C*-algebras with distinct max and min norms on their tensor product, given by Takesaki [Tak64], is for $\mathcal{A} \otimes \mathcal{A}$. We will show an argument that mostly follows [Was90, page 377].

***13.4.28. Corollary.***

*Let $\mathcal{A} = C^*_\lambda(\mathbb{F}_2)$. Then $\mathcal{A} \otimes_{\max} \mathcal{A} \not\cong \mathcal{A} \otimes_{\min} \mathcal{A}$.*

*Proof.* By Theorem 13.3.10 and Corollary 13.4.25, $\mathcal{A} \otimes_{\min} \mathcal{A}$ is simple. On other hand, let $\pi : \mathcal{A} \otimes \mathcal{A} \to \mathcal{B}(\mathcal{H}_{\mathbb{F}_2})$ be given by $\pi(x \otimes y) = xJyJ$. By Proposition 13.3.12, $\pi$ is irreducible. Let $T \in \pi(\mathcal{A} \otimes \mathcal{A})$ be as in Lemma 13.3.9. Choose $f \in C[-1,1]$ such that $f(1) = 1$ and $f|_{[-1,1-\delta]} = 0$. Then $P = f(T) \in \pi(\mathcal{A} \otimes \mathcal{A})$ is a projection. The proof of Lemma 13.3.9 shows that if $\xi \in (\mathbb{C}\delta_e)^\perp$ then $\|\xi - T\xi\| > 0$, so the eigenspace for $1 \in \sigma(T)$ is equal to $\mathbb{C}\delta_e$; that is, $P$ is the rank-one projection onto $\mathbb{C}\delta_e$. As $\pi(\mathcal{A} \otimes \mathcal{A})$ is irreducible, by Corollary 13.2.47 we get that $\mathcal{K}(\mathcal{H}_{\mathbb{F}_2}) \subset \pi(\mathcal{A} \otimes \mathcal{A})$ is a proper ideal. If we define a norm $\alpha$ on $\mathcal{A} \otimes \mathcal{A}$ by $\alpha(x) = \|\pi(x)\|$, we get that $\mathcal{A} \otimes_\alpha \mathcal{A} = \overline{\pi(\mathcal{A} \otimes \mathcal{A})}$ is not simple; so it cannot be isomorphic to $\mathcal{A} \otimes_{\min} \mathcal{A}$. □

It is a consequence of Corollary 13.4.28 that there is no result analog to Lemma 13.4.18 for the min norm.

*13.4.29. Remark* . An alternative version of the proof of Corollary 13.4.28 can be found in [**KR97b**, Example 11.3.14], where instead of the irreducibility they use the fact that $\mathcal{A} \otimes_{\min} \mathcal{A}$ admits a faithful tracial state (this follows from von Neumann algebra considerations, as the von Neumann algebra generated by $\mathcal{A} \otimes \mathcal{A}$ inside $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ is a II$_1$-factor). Then, after constructing $P_e$ as above—their methods differ a bit from those of Lemmas 13.3.8 and 13.3.9 but they build on the same ideas—one notices that $\{\lambda(g)P_e\lambda(g)^*\}$ is an infinite pairwise orthogonal family of projections, and then if $\varphi$ is a tracial state

$$1 = \varphi(I_{\mathcal{H}}) \geq \varphi\Big(\sum_{j=1}^{n} \lambda(g_n)P_e\lambda(g_n)^*\Big) = n\varphi(P_e)$$

for all $n$, so $\varphi$ cannot be faithful.

*13.4.30. Remark*. We used $\mathcal{A} = C^*_\lambda(\mathbb{F}_2)$ in Corollary 13.4.28 because the goal was to show a concrete example, and it was also historically the first example and likely the one with the most direct proof. The same ideas work without change for $\mathcal{A} = C^*_\lambda(\mathbb{F}_n)$, $n \geq 2$. As one might suspect, the key property is the non-amenability of $\mathbb{F}_n$; in fact, if $G$ is a discrete amenable group then $C^*_\lambda(G) \otimes \mathcal{A}$ admits a unique C$^*$-norm for any C$^*$-algebra $\mathcal{A}$. For an entirely different kind of example, Junge–Pisier [**JP95**] proved that $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ admits more than one C$^*$-norm; in fact twenty years later Ozawa–Pisier showed that $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ admits uncountably many non-equivalent C$^*$-norms [**OP16**]. More generally, what Ozawa–Pisier proved is that if $\mathcal{M}, \mathcal{N}$ are von Neumann algebras and $\mathcal{M} \otimes \mathcal{N}$ admits more than one C$^*$-norm, then it admits uncountably many. The problem of whether $C^*(\mathbb{F}_\infty) \otimes C^*(\mathbb{F}_\infty)$ (the full C$^*$-algebra, not the reduced one) admits more than one norm was shown by Kirchberg [**Kir93**] to be equivalent to several conjectures in Operator Algebras, most notably Connes Embedding Problem.

The latter was shown to also be equivalent to many open problems in diverse areas of math, and it was solved in the negative in 2020 via a result in Quantum Complexity by Ji, Natarajan, Vidick, Wright, and Yuen (we refer to [Gol22] for a comprehensive review and references); so $C^*(\mathbb{F}_\infty) \otimes C^*(\mathbb{F}_\infty)$ does admit non-equivalent C*-norms. Interestingly, $C^*(\mathbb{F}_\infty) \otimes \mathcal{B}(\mathcal{H})$ admits a single C*-norm [Kir93, Lemma 2.1 and Proposition 2.2].

*13.4.31. Remark.* Having an example where $\mathcal{A} \otimes_{\max} \mathcal{B} \neq \mathcal{A} \otimes_{\min} \mathcal{B}$, we have the following non-intuitive situation. If $\mathcal{A}_0 \subset \mathcal{A}$ and $\mathcal{B}_0 \subset \mathcal{B}$ are dense subalgebras, the identity map $\mathrm{id} : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}$, with the max norm on the domain and the min norm in the codomain, is a bounded $*$-isomorphism. Its extension to $\mathcal{A} \otimes_{\max} \mathcal{B} \to \mathcal{A} \otimes_{\min} \mathcal{B}$ is a surjective $*$-homomorphism, but it is not injective.

**13.4.8. Nuclearity.** We will take a more or less superficial look at nuclearity, mostly focusing on the equivalence of classic nuclearity with the approximation property. To see more details and/or get a broader perspective, we strongly recommend the first three chapters of [BO08]. We see some benefit to our approach, in that the ideas are condensed to the bare minimum needed to establish the main results. It is still quite a technical tour de force.

We have mentioned many examples where the tensor product of C*-algebras admits more than one C*-norm. There are also many examples of the other kind. So much, that there is even a name for it:

**13.4.32. Definition.**

*A C*-algebra $\mathcal{A}$ is* ***nuclear*** *if $\mathcal{A} \otimes \mathcal{B}$ admits a unique C*-norm for any C*-algebra $\mathcal{B}$.*

*A linear map $\phi : \mathcal{A} \to \mathcal{B}$ is* ***nuclear*** *if there exist nets of contractive completely positive maps $\{\varphi_j\}$ and $\{\psi_j\}$, with $\varphi_j : \mathcal{A} \to M_{n(j)}(\mathbb{C})$ and $\psi_j : M_{n(j)}(\mathbb{C}) \to \mathcal{B}$ such that $\psi_j \circ \varphi_j \to \phi$ pointwise.*

$$\begin{tikzcd} \mathcal{A} \arrow[rr, "\phi"] \arrow[dr, "\varphi_j"'] & & \mathcal{B} \\ & M_{n(j)}(\mathbb{C}) \arrow[ur, "\psi_j"'] & \end{tikzcd}$$

*We say that $\mathcal{A}^{**}$ is* ***semidiscrete*** *if there exist nets of contractive completely positive maps $\{\varphi_j\}$ and $\{\psi_j\}$, with $\varphi_j : \mathcal{A} \to M_{n(j)}(\mathbb{C})$ and $\psi_j : M_{n(j)}(\mathbb{C}) \to \mathcal{B}$ such that $\psi_j \circ \varphi_j \to \mathrm{id}_{\mathcal{A}^{**}}$ pointwise weak*.*

The name *nuclear* was inspired by Grothendieck's amazing work on tensor products of Banach spaces [Gro55], but nuclear C*-algebras are not nuclear as Banach spaces. So far we have seen that $M_n(\mathbb{C})$ is nuclear, and we have mentioned that $C^*_\lambda(G)$ is nuclear when $G$ is discrete and amenable. There seems to be a disconnect between the notions of nuclear C*-algebra and nuclear map, but we will show the close relation momentarily.

We begin with yet another consequence of Takesaki's Theorem.

**13.4.33. Proposition.**

*Let $X$ be a locally compact Hausdorff space and $\mathcal{A}$ a C*-algebra. Given a tensor norm $\alpha$, there is a canonical $*$-isomorphism*

$$C_0(X) \otimes_\alpha \mathcal{A} \simeq C_0(X, \mathcal{A}).$$

*In consequence, any abelian C*-algebra is nuclear.*

*Proof.* Assume first that we are in the unital case. So $X$ is compact and $\mathcal{A}$ is unital. Let $\phi$ be a pure state of $C(X) \otimes_\alpha \mathcal{A}$. Consider the irreducible GNS representation $\pi_\phi$ (Proposition 13.2.41) and write $\pi_\phi|_{C(X)\otimes\mathcal{A}} = \pi_X \times \pi_\mathcal{A}$ as in Proposition 13.4.9. As in the proof of Corollary 13.4.25, the centre of $\pi_X(C(X))$ can only consist of scalar multiples of the identity, which means that $\pi_X = \delta_x$ for some $x \in X$ by Proposition 7.4.6. Now

$$\phi(f \otimes a) = \langle \pi_\phi(f \otimes a)\xi_\phi, \xi_\phi\rangle = \delta_x(f)\,\langle \pi_\mathcal{A}(a)\xi_\phi, \xi_\phi\rangle = (\delta_x \otimes \psi)(f \otimes a),$$

where $\psi \in S(\mathcal{A})$ is the state $\psi(a) = \langle \pi_\mathcal{A}(a)\xi_\phi, \xi_\phi\rangle$. Since $\pi_\mathcal{A}$ is irreducible and it is a GNS representation for $\psi$, we also have that $\psi$ is pure. Corollary 13.4.21 guarantees that $\delta_x \otimes \psi$ is max-bounded. We conclude that $C(X) \otimes_{\max} \mathcal{A}$ and $C(X) \otimes_{\min} \mathcal{A}$ have the same pure states, and hence the same states after taking convex combinations and pointwise limits, as in Corollary 13.2.42. Thus $\|\cdot\|_{\max} = \|\cdot\|_{\min}$ on $C(X) \otimes \mathcal{A}$ by Exercise 11.5.6. In the non-unital case, we can first appeal to Proposition 13.4.13 to reduce to the unital case, so $C_0(X) \otimes \mathcal{A}$ admits a unique C*-norm.

Now we can focus on the isomorphism. The natural map here would be

$$\gamma : \sum_k f_k \otimes a_k \longmapsto \Big(x \longmapsto \sum_k f_k(x)\, a_k\Big).$$

This map is clearly a $*$-homomorphism, **if** it is well-defined. For that, we know by Proposition 13.1.4 that $\sum_k a_k \otimes f_k = 0$ if and only if there exist coefficients $c_{kj}$ such that

$$\sum_n c_{nk} a_n = 0, \qquad \sum_j c_{kj} f_j = f_k \qquad \text{for all } k.$$

Then

$$\sum_k f_k(x)a_k = \sum_k \sum_j c_{kj} f_j(x)a_k = \sum_j f_j(x) \sum_k c_{kj} a_k = 0.$$

So $\gamma$ is well-defined and injective (since the steps can be reversed). So the map $\gamma$ is an injective $*$-homomorphism $C(X) \otimes \mathcal{A} \to C_0(X, \mathcal{A})$. If we now define $\alpha(x) = \|\gamma(x)\|$ for $x \in C_0(X) \otimes \mathcal{A}$, the fact that $\gamma$ is a $*$-monomorphism guarantees that $\alpha$ is a C*-norm. As $C_0(X)$ is nuclear, we have $\|\gamma(x)\| = \|x\|_{\min}$ for al $x \in C_0(X) \otimes \mathcal{A}$.

It remains to be seen that $\gamma$ is surjective. Let $f \in C_c(X, \mathcal{A})$ and fix $\varepsilon > 0$. Since $f$ is continuous with compact support, $f(X)$ is compact. So there exist $a_1, \ldots, a_r \in f(X)$ such that $f(X) \subset \bigcup_k B_\varepsilon(a_k)$. Let $x_k \in X$ with $f(x_k) = a_k$, and $V_k = f^{-1}(B_\varepsilon(a_k))$. These are open sets that cover $X$. Let $h_1, \ldots, h_r$ be a partition of unity relative to $V_1, \ldots, V_r$ (Proposition 2.9.1). That is, $\sum_k h_k = 1$, $0 \leq h_k \leq 1$, continuous, and $\operatorname{supp} h_k \subset V_k$. Let $f_k = f\, h_k$. Then, taking into account that $h_k$ is nonzero only inside $V_k$,

$$\Big\| f(x) - \sum_k h_k(x)a_k \Big\| = \Big\| \sum_k h_k(x) \big(f(x) - f(x_k)\big) \Big\| \leq \varepsilon \Big\| \sum_k h_k(x) \Big\| = \varepsilon.$$

We have shown that image of $\gamma$ is dense. Being isometric, $\gamma$ extends to a $*$-isomorphism $C_0(X) \otimes_{\min} \mathcal{A} \to C_0(X, \mathcal{A})$. As $C_0(X) \otimes_\alpha \mathcal{A} = C_0(X) \otimes_{\min} \mathcal{A}$ by the first part of the proof, we are done. □

A rather straightforward consequence of Proposition 13.4.33 is that for locally compact Hausdorff spaces $X, Y$ we have $C_0(X) \otimes_\alpha C_0(Y) \simeq C_0(X \times Y)$ for any norm $\alpha$ (Exercise 13.4.28).

Here comes a first step towards joining the notion of nuclear C*-algebra and nuclear map.

***13.4.34. Proposition.***

*Let $\mathcal{A}$ be a C*-algebra such that $\mathcal{A}^{**}$ is semidiscrete. Then $\mathrm{id}_{\mathcal{A}}$ is nuclear.*

*Proof.* Assume first that $\mathcal{A}$ is unital. Fix $\varepsilon > 0$, $F \subset \mathcal{A}$ finite, and $K \subset \mathcal{A}^*$ finite (note that $\mathcal{A}^*$ is the predual of $\mathcal{A}^{**}$). By Theorem 12.7.8 we can assume that $\mathcal{A}^{**}$ is a von Neumann algebra with $\mathcal{A}$ $\sigma$-weakly dense in it. By hypothesis there exist contractive completely positive maps $\varphi : \mathcal{A} \to M_n(\mathbb{C})$ and $\psi : M_n(\mathbb{C}) \to \mathcal{A}^{**}$ such that $|\eta(\psi \circ \varphi(a)) - \eta(a)| < \varepsilon$ for all $a \in F$ and all $\eta \in K$. All that one needs to do is to make the codomain to be $\mathcal{A}$ instead of $\mathcal{A}^{**}$, but that's no minor thing.

By Lemma 13.2.62 we may assume that $\varphi$ and $\psi$ are unital. By Choi's Criterion (Theorem 13.2.50) we can identify $\psi$ with the positive matrix $E_\psi = [\psi(E_{kj})] \in M_n(\mathcal{A}^{**})$. We can write $E_\psi = F^*F$ for some $F \in M_n(\mathcal{A}^{**})$. Using that $\mathcal{A}$ is $\sigma$-weakly dense in $\mathcal{A}^{**}$ we get that it is also sot-dense. By Kaplansky

(Theorem 12.3.17) and the fact that convergence in $M_n(\mathcal{A}^{**})$ is achieved entrywise for all vector space topologies, there exists a net $\{E_\ell\} \subset M_n(\mathcal{A})^+$ such that $E_\ell \xrightarrow{\text{sot}} E_\psi$. Let $\psi_\ell : M_n(\mathbb{C}) \to \mathcal{A}$ be the corresponding completely positive maps under the Choi identification. By construction, $\psi_\ell(E_{kj}) \xrightarrow{\text{sot}} \psi(E_{kj})$ for all $k, j$, so $\psi_\ell(A) \xrightarrow{\text{sot}} \psi(A)$ for all $A \in M_n(\mathbb{C})$. Choosing $\ell$ far enough along the net we get a completely positive map $\psi' : M_n(\mathbb{C}) \to \mathcal{A}$ that satisfies $|\eta(\psi' \circ \varphi(a)) - \eta(a)| < \varepsilon$ for all $a \in F$ and all $\eta \in K$. That is, we get nets of contractive completely positive maps $\{\varphi_j\}$ and $\{\psi_j\}$ with $\psi_j \circ \varphi_j(a) \xrightarrow{\text{weak}} a$ for all $a \in \mathcal{A}$. By Proposition 9.1.4 we get a net $\{\theta_F\}$ of linear maps such that each of them is a convex combination of some of the $\psi_j \circ \varphi_j$ (so in particular, they are completely positive) and $\theta_F(a) \to a$; . Then Lemma 13.2.63 gives us $\varphi_F : \mathcal{A} \to M_{n(F)}(\mathbb{C})$ and $\psi : M_{n(F)}(\mathbb{C}) \to \mathcal{A}$, completely positive, such that $\theta_F = \psi_F \circ \varphi_F$ for all $F$. Hence $\mathrm{id}_{\mathcal{A}}$ is nuclear.

Finally, we need to deal with the non-unital case. Since $\mathcal{A}^{**}$ is semidiscrete, there exist nets of contractive completely positive maps $\varphi_j : \mathcal{A}^{**} \to M_{n(j)}(\mathbb{C})$ and $\psi_j : M_{n(j)}(\mathbb{C}) \to \mathcal{A}^{**}$ such that $\psi_j \circ \varphi_j \to \mathrm{id}_{\mathcal{A}^{**}}$ pointwise weak*. By Exercise 12.7.6, $(\tilde{\mathcal{A}})^{**} \simeq \mathcal{A}^{**} \oplus \mathbb{C}$. So we can extend the maps as $\varphi_j'(a, \lambda) = \varphi_j(a) \oplus \lambda \in M_{n(j)+1}(\mathbb{C})$, and $\psi_j : M_{n(j)+1}(\mathbb{C}) \to (\tilde{\mathcal{A}})^{**}$ as $\psi_j'(X) = \psi(X(n(j))) \oplus X_{n(j)+1,n(j)+1}$, where $X(n)$ denotes the matrix obtained from $X$ by erasing the last row and column. As $\tilde{\mathcal{A}}$ is unital, the first part of the proof shows that $\mathrm{id}_{\tilde{\mathcal{A}}}$ is nuclear. And then $\mathrm{id}_{\mathcal{A}}$ is nuclear by Exercise 13.4.26. ☐

The great result about nuclearity, Theorem 13.4.35, follows from [CE76] and [Kir77]. These papers belong to an exhilarating time for operator algebras, where amazing developments happened very rapidly. They form part of the body of work that led to and benefitted from Connes' incredible [Con76].

***13.4.35. Theorem. (Choi–Effros, Kirchberg)***

*Let $\mathcal{A}$ be a C*-algebra. The following statements are equivalent:*

*(i)* $\mathrm{id}_{\mathcal{A}}$ *is nuclear;*

*(ii)* $\mathcal{A}$ *is nuclear.*

*Proof.* (i) $\implies$ (ii) We need to take advantage of the fact that $M_n(\mathbb{C}) \otimes \mathcal{B}$ admits a unique C*-norm (Remark 13.4.12). By hypothesis there exist contractive completely positive maps $\varphi_j : \mathcal{A} \to M_{n(j)}(\mathbb{C})$ and $\psi_j : M_{n(j)}(\mathbb{C}) \to \mathcal{A}$ such that $\psi_j \circ \varphi_j(a) \to a$ for all $a \in \mathcal{A}$. Let $\mathcal{B}$ be a C*-algebra. We use

Lemma 13.4.17 to construct $\psi_j \otimes_{\max} \mathrm{id}_{\mathcal{B}} : M_{n(j)}(\mathbb{C}) \otimes \mathcal{B} \to \mathcal{A} \otimes_{\max} \mathcal{B}$. Because $M_{n(j)}(\mathbb{C})$ is nuclear,

$$\|(\psi_j \otimes_{\max} \mathrm{id}_{\mathcal{B}})(z)\|_{\max} \leq \|z\|_{\max} = \|z\|_{\min}.$$

By Lemma 13.4.26 we have completely positive maps $\varphi_j \otimes_{\min} \mathrm{id}_{\mathcal{B}} : \mathcal{A} \otimes_{\min} \mathcal{B} \to M_{n(j)}(\mathbb{C}) \otimes \mathcal{B}$. The map $\Theta_j : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{A} \otimes_{\max} \mathcal{B}$, given by $\Theta_j = (\psi_j \otimes_{\max} \mathrm{id}_{\mathcal{B}}) \circ (\varphi_j \otimes_{\min} \mathrm{id}_{\mathcal{B}})$, is bounded (contractive, in fact). And, for $x = \sum_k a_k \otimes b_k \in \mathcal{A} \otimes \mathcal{B}$,

$$\begin{aligned}\|x - \Theta_j(x)\|_{\max} &= \left\| \sum_k (a_k - \psi_j(\varphi_j(a_k))) \otimes b_k \right\|_{\max} \\ &\leq \sum_k \|a_k - \psi_j(\varphi_j(a_k))\| \, \|b_k\| \xrightarrow[j]{} 0.\end{aligned}$$

Then

$$\|x\|_{\max} = \lim_j \|\Theta_j(x)\|_{\max} \leq \|x\|_{\min}.$$

Hence the max and min norms agree on $\mathcal{A} \otimes \mathcal{B}$ and $\mathcal{A}$ is nuclear.

(ii) $\Longrightarrow$ (i) Because of Corollary 13.4.27 we may assume without loss of generality that $\mathcal{A}$ is unital. Let $\gamma : \mathcal{A} \to \mathcal{A}^{**}$ be the canonical inclusion. For notational purposes, let $\mathcal{M} = \mathcal{A}^{**} \subset \mathcal{B}(\mathcal{H})$. Consider the representation $\gamma \times \mathrm{id}_{\mathcal{M}'} : \mathcal{A} \otimes \mathcal{M}' \to \mathcal{B}(\mathcal{H})$ induced by $(\gamma \times \mathrm{id}_{\mathcal{M}'})(a \otimes T) = \gamma(a)T$. Then

$$\|(\gamma \times \mathrm{id}_{\mathcal{M}'})(X)\| \leq \|X\|_{\max} = \|X\|_{\min}, \qquad X \in \mathcal{A} \otimes \mathcal{M}',$$

the inequality by definition of the max norm, and the equality by the nuclearity of $\mathcal{A}$. Fix $\varepsilon > 0$, $F \subset \mathcal{A}$ finite, and $\mathcal{S} = \{\eta_1, \ldots, \eta_n\} \subset \mathcal{M}_*^+$ finite. Put $\eta = \frac{1}{n} \sum_k \eta_k$. Then $\eta_k \leq n\eta$ for all $k$, and by Theorem 13.2.38 there exists $T_k \in \pi_\eta(\mathcal{M})'$ such that $\eta_k(X) = \langle \pi_\eta(X) T_k \xi_\eta, \xi_\eta \rangle$ (Theorem 13.2.38 is phrased for completely positive maps, but we know that positive linear functionals are completely positive by Proposition 13.2.22 or Proposition 13.2.12, and the GNS representation is the Stinespring Dilation in such case). We claim that $(\pi_\eta \circ \gamma) \times \mathrm{id}_{\pi_\eta(\mathcal{M})'} : \mathcal{A} \otimes \pi_\eta(\mathcal{M})' \to \mathcal{B}(\mathcal{H}_\eta)$ is min-continuous—we leave the proof till the end. Let $\mu : \mathcal{A} \otimes \pi_\eta(\mathcal{M})' \to \mathbb{C}$ be linearly induced by

$$\mu(a \otimes T) = \langle \pi_\eta(\gamma(a)) T \xi_\eta, \xi_\eta \rangle.$$

The min-continuity claimed above guarantees that $\mu$ extends to a state on $\mathcal{A} \otimes_{\min} \pi_\eta(\mathcal{M})'$. We have in particular that $\mu(a \otimes T_k) = \eta_k(\gamma(a))$ for all $a \in \mathcal{A}$. Let $\mathcal{A}_F = C^*(F) \subset \mathcal{A}$ and $\mathcal{B} = C^*(T_1, \ldots, T_n) \subset \pi_\eta(\mathcal{M})'$; being finitely generated, $\mathcal{A}_F$ and $\mathcal{B}$ are both separable. Let $\rho : \mathcal{A}_F \to \mathcal{B}(\mathcal{K}_0)$ be a faithful representation. Then $\tilde{\rho} : \mathcal{A} \to \mathcal{B}(\mathcal{K}_0 \otimes \ell^2(\mathbb{N}))$ given by $\tilde{\rho}(a) = \rho(a) \otimes I_{\ell^2(\mathbb{N})}$ is a faithful representation such that $\tilde{\rho}(\mathcal{A})$ contains no compact operators. Similarly, $\tilde{\mathcal{B}} = \mathcal{B} \otimes I_{\ell^2(\mathbb{N})}$ contains no compact operators.

With $\mathcal{K} = \mathcal{K}_0 \mathbin{\overline{\otimes}} \ell^2(\mathbb{N})$ we consider the faithful representation

$$\tilde{\rho} \otimes_{\min} (\mathrm{id}_{\mathcal{B}} \otimes I_{\ell^2(\mathbb{N})}) : \mathcal{A} \otimes_{\min} \mathcal{B} \to \mathcal{B}(\mathcal{K} \mathbin{\overline{\otimes}} (\mathcal{H}_\eta \mathbin{\overline{\otimes}} \ell^2(\mathbb{N})))$$

(this exists by Lemma 13.4.26). Let $\tilde\mu \in S(\tilde\rho(\mathcal{A}) \otimes_{\min} (\mathcal{B} \otimes_{\min} I_{\ell^2(\mathbb{N})}))$ be the state given by

$$\tilde\mu(x) = \mu\big((\tilde\rho \otimes_{\min} (\mathrm{id}_{\mathcal{B}} \otimes I_{\ell^2(\mathbb{N})})^{-1}(x)\big).$$

We have $\mu(a \otimes T_k) = \tilde\mu(\tilde\rho(a) \otimes (T_k \otimes I_{\ell^2(\mathbb{N})}))$.

By Glimm's Lemma (Theorem 13.2.49) we can write $\tilde\mu$ as a limit of vector states. Further approximating the vectors by vectors on the algebraic tensor product, there exist orthonormal $\{\xi_1, \dots, \xi_r\} \subset \mathcal{K}$ and $\{S_{t,g} : 1 \le t \le r,\ 1 \le g \le m\} \subset \mathcal{M}$ such that (using the notation $\hat S = \pi_\eta(S)\xi_\eta$)

$$\left|\mu(a \otimes T_k) - \Big\langle \sum_{g=1}^m (\tilde\rho(a) \otimes T_k)\Big(\sum_{t=1}^r \xi_t \otimes \hat S_{t,g}\Big), \sum_{t=1}^r \xi_t \otimes \hat S_{t,g}\Big\rangle\right| < \varepsilon$$

for all $a \in F$ and $k = 1, \dots, n$. This requires a bit of unpacking, since the third component of the tensor product is gone. What happens is that the second term in the difference above looks like (writing $\tilde X = \tilde\rho(a) \otimes (T_k \otimes I_{\ell^2(\mathbb{N})})$)

$$\begin{aligned}
&\Big\langle \tilde X \Big(\sum_{t=1}^r \sum_{g=1}^m \xi_t \otimes \hat S_{t,g} \otimes \delta_{e_g}\Big), \sum_{t=1}^r \sum_{j=1}^m \xi_t \otimes T_k \hat S_{t,g} \otimes \delta_{e_g}\Big\rangle \\
&\qquad = \sum_{t,s=1}^r \sum_{g,h=1}^m \Big\langle \tilde\rho(a)\xi_t \otimes T_k \hat S_{t,g}, \xi_s \otimes T_k \hat S_{s,h}\Big\rangle \langle \delta_{e_g}, \delta_{e_h}\rangle \\
&\qquad = \sum_{t,s=1}^r \sum_{g=1}^m \Big\langle \tilde\rho(a)\xi_t \otimes T_k \hat S_{t,g}, \xi_s \otimes T_k \hat S_{s,g}\Big\rangle.
\end{aligned}$$

Now let $P \in \mathcal{B}(\mathcal{K})$ be the orthogonal projection onto $\mathrm{span}\{\xi_1, \dots, \xi_r\}$ and let $\varphi : \mathcal{A} \to P\mathcal{B}(\mathcal{K})P$ be the completely positive map given by $\varphi(a) = P\tilde\rho(a)P$. When we identify $P\mathcal{B}(\mathcal{K})P$ with $M_r(\mathbb{C})$ and $\{\xi_j\}$ with the canonical basis in $\mathbb{C}^r$, we have

$$\varphi(a) = \sum_{k,j=1}^r \langle \tilde\rho(a)\xi_j, \xi_k\rangle E_{kj}.$$

We also define $\psi : M_r(\mathbb{C}) \to \mathcal{M}$ by $\psi(E_{kj}) = \sum_{g=1}^m S_{k,g}^* S_{j,g}$. This is completely positive by Choi's Criterion (Theorem 13.2.50) since

$$\begin{aligned}
\psi^{(r)}\Big(\sum_{k,j=1}^r E_{kj} \otimes E_{kj}\Big) &= \sum_{g=1}^m \sum_{k,j=1}^r S_{k,g}^* S_{j,g} \otimes E_{k,j} \\
&= \sum_{g=1}^m \sum_{k,j} (S_{k,g} \otimes E_{1,k})^* (S_{j,g} \otimes E_{1,j}) \\
&= \sum_{g=1}^m \Big(\sum_{k=1}^r S_{k,g} \otimes E_{1,k}\Big)^* \Big(\sum_{k=1}^r S_{k,g} \otimes E_{1,k}\Big) \ge 0.
\end{aligned}$$

Since $T_k \in \pi_\eta(\mathcal{M})'$,

$$\eta_k(S_{h,g}^* S_{j,g}) = \langle \pi_\eta(S_{h,g}^* S_{j,g}) T_k \xi_\eta, \xi_\eta \rangle = \langle T_k \hat{S}_{j,g}, \hat{S}_{h,g} \rangle.$$

Given $a \in F$ and $k$, with $r = |\eta_k(\gamma(a)) - \eta_k(\psi \circ \varphi(a))|$,

$$\begin{aligned} r &= \left| \mu(a \otimes T_k) - \sum_{g=1}^{m} \sum_{h,j=1}^{r} \langle \tilde{\rho}(a)\xi_j, \xi_h \rangle \, \eta_k(S_{h,g}^* S_{j,g}) \right| \\ &= \left| \mu(a \otimes T_k) - \sum_{g=1}^{m} \sum_{h,j=1}^{r} \langle \tilde{\rho}(a)\xi_j, \xi_h \rangle \, \langle T_k \hat{S}_{j,g}, \hat{S}_{h,g} \rangle \right| \\ &= \left| \mu(a \otimes T_k) - \Big\langle \sum_{g=1}^{m} (\tilde{\rho}(a) \otimes T_k) \Big( \sum_{t=1}^{r} \xi_t \otimes \hat{S}_{t,g} \Big), \sum_{t=1}^{r} \xi_t \otimes \hat{S}_{t,g} \Big\rangle \right| \\ &< \varepsilon. \end{aligned}$$

As we can do this for all $\varepsilon$, $F$, and $\mathcal{S}$, it means that we can construct nets $\{\varphi_\ell\}$ with $\varphi_\ell : \mathcal{A} \to M_{n(\ell)}(\mathbb{C})$ and $\{\psi_\ell\}$ with $\psi_\ell : M_{n(\ell)}(\mathbb{C}) \to \mathcal{M}$, all completely positive, and such that $\psi_\ell \circ \varphi_\ell \to \gamma$ pointwise $\sigma$-weak. We have shown that $\mathcal{A}^{**}$ is semidiscrete, and therefore $\mathrm{id}_{\mathcal{A}}$ is nuclear by Proposition 13.4.34.

To finish the proof, it remains to show the claim that $(\pi_\eta \circ \gamma) \times \mathrm{id}_{\pi_\eta(\mathcal{M})'} : \mathcal{A} \otimes \pi_\eta(\mathcal{M})' \to \mathcal{B}(\mathcal{H}_\eta)$ is min-continuous. Because $\pi_\eta$ is normal, its kernel is wot-closed by Proposition 12.6.11 and so $\ker \pi_\eta = Z\mathcal{M}$ for some $Z \in \mathcal{Z}(\mathcal{M})$ by Corollary 12.3.12. The representation $\pi_\eta$ can be seen as $\pi'_\eta(I_{\mathcal{M}} - Z)$, where $\pi'_\eta$ is the faithful normal restriction of $\pi_\eta$ to $(I_{\mathcal{M}} - Z)\mathcal{M}$. As multiplication by a projection is $\sigma$-weakly continuous, we may assume without loss of generality that $\pi_\eta$ is faithful. Then Corollary 12.6.14 and Corollary 12.6.16 imply that $\pi_\eta$ is unitarily equivalent to a compression by a commuting projection of an amplification of the identity representation. Both unitary equivalence and compression by a projection are min-continuous (as they are continuous for any norm), so we may assume that $\pi_\eta(T) = T \otimes I_{\mathcal{K}} \in \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ for some Hilbert space $\mathcal{K}$ (see Exercise 13.4.12). In this situation we have by Exercise 13.4.27

$$(\mathcal{M} \otimes_{\min} I_{\mathcal{K}})' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}),$$

where $\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$ is the von Neumann algebra generated by $\mathcal{M}' \otimes \mathcal{B}(\mathcal{K})$ in $\mathcal{B}(\mathcal{H} \otimes \mathcal{K})$. We have thus reduced the problem to showing that

$$(\gamma \otimes I_{\mathcal{K}}) \times \mathrm{id}_{\mathcal{M}' \overline{\otimes} \mathcal{B}(\mathcal{K})} : \mathcal{A} \otimes (\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})) \to \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$$

is min-continuous. Suppose first that $\dim \mathcal{K} < \infty$. In that case $\mathcal{M}' \otimes \mathcal{B}(\mathcal{K})$ is already a von Neumann algebra. Our map sends $a \otimes T \otimes S$, for $a \in \mathcal{A}$, $T \in \mathcal{M}'$ and $S \in \mathcal{B}(\mathcal{K})$, to $(\gamma(a) \otimes I_{\mathcal{K}})(T \otimes S) = \gamma(a)T \otimes S$; this can be seen as the image of $(\gamma \times \mathrm{id}_{\mathcal{M}'}) \otimes \mathrm{id}_{\mathcal{B}(\mathcal{K})} : (\mathcal{A} \otimes \mathcal{M}') \otimes \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$. By Lemma 13.4.18 we have that $\gamma \times \mathrm{id}_{\mathcal{M}'}$ is max-bounded; but $\mathcal{A}$ is nuclear, so the min and max norms agree on $\mathcal{A} \otimes \mathcal{M}'$, and thus $\gamma \times \mathrm{id}_{\mathcal{M}'}$ is min-bounded.

And $(\gamma\times\mathrm{id}_{\mathcal{M}'})\otimes\mathrm{id}_{\mathcal{B}(\mathcal{K})}$ can be seen as the $\dim\mathcal{K}$-amplification of $(\gamma\times\mathrm{id}_{\mathcal{M}'})$, which has the same norm as $(\gamma\times\mathrm{id}_{\mathcal{M}'})$ since the latter is completely positive. Finally, we let $\{P_j\}$ be an increasing net of finite-rank projections such that $P_j \xrightarrow{\text{sot}} I_{\mathcal{K}}$. Let

$$X = \sum_k a_k \otimes (T_k \otimes S_k) \in \mathcal{A} \otimes (\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})).$$

Given $\xi\otimes\eta \in \mathcal{H}\otimes\mathcal{K}$ and $\Theta = (\gamma\otimes I_{\mathcal{K}})\times\mathrm{id}_{\mathcal{M}'\overline{\otimes}\mathcal{B}(\mathcal{K})})$ we have (since the sum is finite)

$$\begin{aligned}\Theta(X)\,(\xi\otimes\eta) &= \sum_k \gamma(a_k)T_k\xi\otimes S_k\eta = \lim_j \sum_k \gamma(a_k)T_k\xi\otimes P_jS_kP_j\eta \\ &= \lim_j \big((\gamma\otimes I_{P_j\mathcal{K}})\times\mathrm{id}_{\mathcal{M}'\overline{\otimes}\mathcal{B}(P_j\mathcal{K})}\big)\big((I_{\mathcal{H}}\otimes P_j)X(I_{\mathcal{H}}\otimes P_j)\big)(\xi\otimes\eta).\end{aligned}$$

By taking sums the equality applies to all of $\mathcal{H}\otimes\mathcal{K}$. Therefore, for any unit vector $\tilde{\xi}\in\mathcal{H}\otimes\mathcal{K}$

$$\begin{aligned}\|\Theta(X)\tilde{\xi}\| &= \lim_j \big\|\big((\gamma\otimes I_{P_j\mathcal{K}})\times\mathrm{id}_{\mathcal{M}'\overline{\otimes}\mathcal{B}(P_j\mathcal{K})}\big)\big((I_{\mathcal{H}}\otimes P_j)X(I_{\mathcal{H}}\otimes P_j)\big)\,\tilde{\xi}\big\| \\ &\le \limsup_j \big\|\big((\gamma\otimes I_{P_j\mathcal{K}})\times\mathrm{id}_{\mathcal{M}'\overline{\otimes}\mathcal{B}(P_j\mathcal{K})}\big)\big\|_{\min}\|(I_{\mathcal{H}}\otimes P_j)X(I_{\mathcal{H}}\otimes P_j)\|_{\min} \\ &\le \limsup_j \|\gamma\times\mathrm{id}_{\mathcal{M}'}\|_{\min}\,\|I_{\mathcal{H}}\otimes P_j\|_{\min}^2\,\|X\|_{\min} \\ &\le \limsup_j \|\gamma\times\mathrm{id}_{\mathcal{M}'}\|_{\min}\,\|P_j\|^2\,\|X\|_{\min} \\ &= \|\gamma\times\mathrm{id}_{\mathcal{M}'}\|_{\min}\,\|X\|_{\min}.\end{aligned}$$

Hence $\|\Theta X\| \le \|\gamma\times\mathrm{id}_{\mathcal{M}'}\|_{\min}\,\|X\|_{\min}$, showing that $\Theta$ is min-continuous. □

**13.4.9. Exercises.**

(13.4.1) Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces with orthonormal bases $\{\xi_j\}$ and $\{\eta_k\}$, respectively. Show that $\{\xi_j\otimes\eta_k\}$ is an orthonormal basis for $\mathcal{H}\mathbin{\overline{\otimes}}\mathcal{K}$.

(13.4.2) Prove the isomorphisms (13.20).

(13.4.3) Show that the tensor product of operators obeys the same arithmetic rules as the elementary tensors of vectors, as in Proposition 13.1.2.

(13.4.4) Let $\mathcal{M}\subset\mathcal{B}(\mathcal{K})$ be a von Neumann algebra and $\mathcal{H}$ a Hilbert space. Fix an orthonormal basis $\{\eta_j\}_{j\in J}$ for $\mathcal{H}$ and consider the associated matrix units $\{E_{kj}\}$. Show that for each $\tilde{T}\in\mathcal{M}\mathbin{\overline{\otimes}}\mathcal{B}(\mathcal{H})$ there exist unique operators $\{T_{kj}\}\subset\mathcal{M}$ such that

$$\tilde{T} = \sum_{k,j} T_{kj}\otimes E_{kj},$$

where the series converges sot.

(13.4.5) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate von Neumann algebra and consider matrix units $\{E_{kj}\}_{k,j\in J} \subset \mathcal{M}$ such that $\sum_k E_{kk} = I_{\mathcal{M}}$ (with the series converging sot). Fix $j_0 \in J$ and let $P = E_{j_0,j_0}$. Show that there exists a Hilbert space $\mathcal{K}$ such that $\mathcal{M} \simeq P\mathcal{M}P \,\overline{\otimes}\, \mathcal{B}(\mathcal{K})$ (we take as assumed knowledge that $P\mathcal{M}P \subset \mathcal{B}(P\mathcal{H})$ is a von Neumann algebra; this will be proven in Proposition 14.1.1).

(13.4.6) Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Show that

$$\mathcal{K}(\mathcal{H}_1 \,\overline{\otimes}\, \mathcal{H}_2) = \overline{\mathcal{K}(\mathcal{H}_1) \otimes \mathcal{K}(\mathcal{H}_2)}.$$

(13.4.7) Let $S \in \mathcal{T}(\mathcal{H})$, $T \in \mathcal{T}(\mathcal{K})$. Show that $S \otimes T \in \mathcal{T}(\mathcal{H} \,\overline{\otimes}\, \mathcal{K})$.

(13.4.8) Let $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ be Hilbert spaces. Show that

$$(\mathcal{H}_1 \,\overline{\otimes}\, \mathcal{H}_2) \,\overline{\otimes}\, \mathcal{H}_3 \simeq \mathcal{H}_1 \,\overline{\otimes}\, (\mathcal{H}_2 \,\overline{\otimes}\, \mathcal{H}_3)$$

canonically.

(13.4.9) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space, and $n \in \mathbb{N}$. Show that $M_n(\mathbb{C}) \otimes \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C}) \,\overline{\otimes}\, \mathcal{B}(\mathcal{H}) \simeq \mathcal{B}(\mathcal{H})$ as von Neumann algebras.

(13.4.10) Let $\mathcal{M}_k \subset \mathcal{B}(\mathcal{H}_k)$, $k = 1, 2, 3$, be von Neumann algebras. Show that there is a canonical isomorphism

$$(\mathcal{M}_1 \,\overline{\otimes}\, \mathcal{M}_2) \,\overline{\otimes}\, \mathcal{M}_3 \simeq \mathcal{M}_1 \,\overline{\otimes}\, (\mathcal{M}_2 \,\overline{\otimes}\, \mathcal{M}_3).$$

(13.4.11) Let $\mathcal{H}, \mathcal{K}$ be separable infinite-dimensional Hilbert spaces. Show that the subalgebras

$$\mathcal{K}(\mathcal{H}) \otimes_{\min} \mathcal{K}(\mathcal{K}), \qquad \mathcal{K}(\mathcal{H}) \otimes_{\min} \mathcal{B}(\mathcal{K}),$$

and

$$\mathcal{B}(\mathcal{H}) \otimes_{\min} \mathcal{K}(\mathcal{K})$$

are three distinct ideals of $\mathcal{B}(\mathcal{H}) \otimes_{\min} \mathcal{B}(\mathcal{K})$.

(13.4.12) Let $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. In Definition 12.6.15 we considered the **amplification** of $\pi$ given by $\tilde{\pi} : \mathcal{A} \to \bigoplus_{k\in K} \mathcal{H}$. Show that there is a unitary $U : \bigoplus_{k\in K} \mathcal{H} \to \mathcal{H} \,\overline{\otimes}\, \mathcal{K}$, where $\mathcal{K}$ is a Hilbert space with $\dim \mathcal{K} = |K|$, such that $\tilde{\pi} = U^*(\pi \otimes I_{\mathcal{K}})U$.

(13.4.13) Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces and $\{P_j\}_{j\in J} \subset \mathcal{B}(\mathcal{H})$, $\{Q_k\}_{k\in K} \subset \mathcal{B}(\mathcal{K})$ increasing nets of projections. Show that $\{P_j \otimes Q_k\} \subset \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$ is an increasing net of projections. Show that if $P_j \xrightarrow{\text{sot}} I_{\mathcal{H}}$ and $Q_k \xrightarrow{\text{sot}} I_{\mathcal{K}}$ then $P_j \otimes Q_k \xrightarrow{\text{sot}} I_{\mathcal{H}\overline{\otimes}\mathcal{K}}$.

(13.4.14) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras. Use Proposition 13.1.4 to show that the product and involution are well-defined on $\mathcal{A} \otimes \mathcal{B}$.

(13.4.15) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras. Show that the product on $\mathcal{A}\otimes\mathcal{B}$ is well-defined (use Theorem 13.1.6 and Corollary 13.1.8).

(13.4.16) Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be unital C*-algebras and $\pi : \mathcal{A}\otimes\mathcal{B} \to \mathcal{C}$ a $*$-homomorphism. Show that there exist $*$-homomorphisms $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{C}$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{C}$, with commuting ranges, such that $\pi = \pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$.

(13.4.17) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a representation. Fix $a \in \mathcal{A}$. Show that the map $\rho : b \longmapsto \pi(a \otimes b)$ is linear and bounded. *(Hint: for the bounded part, assume that $a \geq 0$ so that $\varphi \circ \rho$ is positive for any state $\varphi$, and use the Closed Graph Theorem)*

(13.4.18) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation. Show that there exist non-degenerate representations $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}(\mathcal{H})$, with commuting ranges, such that $\pi = \pi_{\mathcal{A}} \times \pi_{\mathcal{B}}$. When $\pi$ is faithful, so are $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$.

*(As opposed to Exercise 13.4.16 the algebras are not required to be unital, so a different method is required; use the non-degeneracy to define $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ on a dense subspace of $\mathcal{H}$, and use Exercise 13.4.17 when needed)*

(13.4.19) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B}(\mathcal{H})$ a non-degenerate representation. Show that if $\{e_\ell\}$ and $\{f_t\}$ are approximate units for $\mathcal{B}$ and $\mathcal{A}$ respectively, the representations $\pi_{\mathcal{A}}$ and $\pi_{\mathcal{B}}$ of Exercise 13.4.18 satisfy

$$\pi_{\mathcal{A}}(a) = \lim_\ell \pi(a \otimes e_\ell), \qquad\qquad \pi_{\mathcal{B}}(b) = \lim_t \pi(f_t \otimes b),$$

where the limits are sot.

(13.4.20) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras and $\pi_{\mathcal{A}} : \mathcal{A} \to \mathcal{H}_{\mathcal{A}}$ and $\pi_{\mathcal{B}} : \mathcal{B} \to \mathcal{H}_{\mathcal{B}}$ cyclic representations. Show that $\pi_{\mathcal{A}} \otimes \pi_{\mathcal{B}}$ is cyclic.

(13.4.21) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras with $\mathcal{A}$ non-unital, $\gamma$ a C*-norm on $\tilde{\mathcal{A}} \otimes \mathcal{B}$, and $\psi \in S(\tilde{\mathcal{A}})$ the unique state with $\ker \psi = \mathcal{A}$. Show that for the map $\psi \otimes_\gamma \mathrm{id}_{\mathcal{B}} : \tilde{\mathcal{A}} \otimes_\gamma \mathcal{B} \to \mathbb{C} \otimes \mathcal{B}$ we have $\ker(\psi \otimes_\gamma \mathrm{id}_{\mathcal{B}}) = \mathcal{A} \otimes_\gamma \mathcal{B}$.

(13.4.22) Let $\mathcal{A}, \mathcal{B}$ be C*-algebras, $x \in \mathcal{A} \otimes_{\min} \mathcal{B}$. Show that if $(\varphi \otimes \psi)(x^*x) = 0$ for all $\varphi \in S(\mathcal{A})$ and $\psi \in S(\mathcal{B})$, then $x = 0$.

(13.4.23) Let $\mathcal{A} \subset \mathcal{B}(\mathcal{H}_{\mathcal{A}})$, $\mathcal{B} \subset \mathcal{B}(\mathcal{H}_{\mathcal{B}})$ be separable C*-algebras. Show that if $\mathcal{A} \otimes_{\min} \mathcal{B} \subset \mathcal{B}(\mathcal{H}_{\mathcal{A}} \overline{\otimes} \mathcal{H}_{\mathcal{B}})$ contains a nonzero compact operator, then $\mathcal{A}$ contains a nonzero compact operator.

(13.4.24) Let $X$ be set, $\mathcal{A}$ a C*-algebra, and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a representation. Let $\delta : c_0(X) \to \mathcal{B}(\ell^2(X))$ be given by $\delta(f) = M_f$. Given $Z = \sum_k f_k \otimes a_k \in c_0(X) \otimes \mathcal{A}$, show that

$$\Big\|(\delta \otimes \pi)(Z)\Big\|_{\mathcal{B}(\ell^2(X)\overline{\otimes}\mathcal{H})} = \sup\Big\{\Big\|(\delta_x \otimes \pi)(Z)\Big\|_{\mathcal{B}(\mathcal{H})} : x \in X\Big\}.$$

(13.4.25) Let $\mathcal{B} \subset \mathcal{A}$ be C*-algebras, with $\mathcal{B}$ hereditary. Show that if $\mathrm{id}_{\mathcal{A}}$ is nuclear, then so is $\mathrm{id}_{\mathcal{B}}$. *(Hint: use an approximate identity for $\mathcal{B}$)*

(13.4.26) Let $\mathcal{A}$ be a C*-algebra. Show that $\mathrm{id}_{\mathcal{A}}$ is nuclear if and only if $\mathrm{id}_{\tilde{\mathcal{A}}}$ is nuclear.

(13.4.27) Let $\mathcal{M}$ be a von Neumann algebra. Show that

$$(\mathcal{M} \otimes_{\min} I_{\mathcal{K}})' = \mathcal{M}' \,\overline{\otimes} \mathcal{B}(\mathcal{K}),$$

where $\mathcal{M}'\,\overline{\otimes}\mathcal{B}(\mathcal{K})$ is the von Neumann algebra generated by $\mathcal{M}'\otimes\mathcal{B}(\mathcal{K})$ in $\mathcal{B}(\mathcal{H}\,\overline{\otimes}\,\mathcal{K})$. Conclude that $\mathcal{Z}(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{K}))=\mathcal{Z}(\mathcal{M})\,\overline{\otimes}\,I_{\mathcal{K}}$.

(13.4.28) Let $X,Y$ be locally compact Hausdorff spaces. Show that

$$C_0(X)\otimes_\alpha C_0(Y)\simeq C_0(X\times Y)$$

for any C*-norm $\alpha$.

## 13.5. Crossed Products

It is a more or less common situation to have a group $G$ acting by automorphisms on a C*-algebra $\mathcal{A}$. For instance $G=\mathbb{R}$ acts on $\mathcal{A}=C_0(\mathbb{R})$ by $(r\cdot f)(t)=f(t+r)$. The idea behind a crossed product—similar in spirit to the semidirect product for groups—is to construct a larger algebra where the automorphisms induced by $G$ are implemented by unitaries.

*13.5.1. Remark* . Given a C*-algebra $\mathcal{A}$, a group $G$ and a homomorphism $\alpha:G\to\mathrm{Aut}(\mathcal{A})$, we can always achieve the goal of having $\mathcal{A}$ living in a larger algebra where the automorphisms induced by $G$ are inner. Fix a non-degenerate representation $\pi:\mathcal{A}\to\mathcal{B}(\mathcal{H})$ and form $\tilde{\mathcal{H}}=\mathcal{H}\,\overline{\otimes}\,\ell^2(G)$. Then we define $\tilde\pi:\mathcal{A}\to\mathcal{B}(\tilde{\mathcal{H}})$ induced by

$$\tilde\pi(a)(\xi\otimes\delta_g)=\pi(\alpha_g^{-1}(a))\xi\otimes\delta_g.$$

With $\lambda:G\to\mathcal{B}(\ell^2(G))$ the left regular representation, the automorphism $\alpha_g$ is induced on $\tilde\pi(\mathcal{A})$ by the unitaries $I_{\mathcal{H}}\otimes\lambda_g$. Indeed,

$$\begin{aligned}\big((I_{\mathcal{H}}\otimes\lambda_g)\tilde\pi(a)(I_{\mathcal{H}}\otimes\lambda_g)^*\big)(\xi\otimes\delta_h)&=(I_{\mathcal{H}}\otimes\lambda_g)\tilde\pi(a)(I_{\mathcal{H}}\otimes\lambda_{g^{-1}})\,(\xi\otimes\delta_h)\\&=(I_{\mathcal{H}}\otimes\lambda_g)\tilde\pi(a)\,(\xi\otimes\delta_{g^{-1}h})\\&=(I_{\mathcal{H}}\otimes\lambda_g)\,(\pi(\alpha_{g^{-1}h}^{-1}(a))\xi\otimes\delta_{g^{-1}h})\\&=(I_{\mathcal{H}}\otimes\lambda_g)\,(\pi(\alpha_{h^{-1}g}(a))\xi\otimes\delta_{g^{-1}h})\\&=(\pi(\alpha_{h^{-1}}\alpha_g(a)\xi\otimes\delta_h)\\&=\tilde\pi(\alpha_g(a))\,(\xi\otimes\delta_h).\end{aligned}$$

Hence

$$(I_{\mathcal{H}}\otimes\lambda_g)\tilde\pi(a)(I_{\mathcal{H}}\otimes\lambda_g)^*=\tilde\pi(\alpha_g(a)).$$

**13.5.2. Definition.**

*A* ***C*-dynamical system*** *is* $(\mathcal{A}, G, \alpha)$ *where* $G$ *is a discrete group,* $\mathcal{A}$ *is a* $C^*$*-algebra, and* $\alpha : G \to \operatorname{Aut}(\mathcal{A})$ *is a group homomorphism. A* ***covariant representation*** *of* $(\mathcal{A}, G, \alpha)$ *is* $(\pi, U)$*, where* $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ *and* $U : G \to \mathcal{B}(\mathcal{H})$ *are representations with* $U_g$ *a unitary and such that*

$$\pi(\alpha_g(a)) = U_g \pi(a) U_g^*.$$

*A covariant representation of the form* $(\tilde{\pi}, I_{\mathcal{H}} \otimes \lambda)$ *as in Remark 13.5.1 is called a* ***regular representation****.*

*The space* $\mathcal{A} \cdot G = \operatorname{span}\{a \cdot g : \ a \in \mathcal{A}, \ g \in G\}$ *is made a* $*$*-algebra with the product induced by* $(a \cdot g)(b \cdot h) = a\alpha_g(b) \cdot gh$ *and the adjoint* $(a \cdot g)^* = \alpha_g^{-1}(a^*) \cdot g^{-1}$*.*

*The* ***C*-algebra crossed product*** $\mathcal{A} \rtimes_\alpha G$ *is the completion of* $\mathcal{A} \cdot G$ *with respect to the norm*

$$\Big\| \sum_{g \in F} a_g \cdot g \Big\| = \sup \Big\{ \Big\| \sum_{g \in F} \pi(a_g)\, U_g \Big\| : \ (\pi, U) \text{ covariant} \Big\}$$

*for each finite* $F \subset G$*. The* ***reduced C*-algebra crossed product*** $\mathcal{A} \rtimes_\alpha^r G$ *is the completion of* $\mathcal{A} \cdot G$ *with respect to the norm obtained as above but considering only the regular representations.*

When we think of the elements of the group as unitaries that implement the automorphisms $\alpha_g$, the definition of product and adjoint above are the natural ones. Indeed, we would have (where the goal is to exchange $g$ and $b$)

$$(a \cdot g)(b \cdot h) = agbh = a(gbg^{-1})gh = a\alpha_g(b) \cdot gh \tag{13.25}$$

and

$$(a \cdot g)^* = g^{-1}a^* = g^{-1}a^* g\, g^{-1} = \alpha_{g^{-1}}(a^*) \cdot g^{-1} = \alpha_g^{-1}(a^*) \cdot g^{-1}. \tag{13.26}$$

The name *covariant representation* is coherent in the sense that each covariant representation $(\pi, U)$ induces a representation $\pi \rtimes U : \mathcal{A} \cdot G \to \mathcal{B}(\mathcal{H})$ by

$$(\pi \rtimes U)\Big( \sum_{g \in F} a_g \cdot g \Big) = \sum_{g \in F} \pi(a_g) U_g$$

for each $F \subset G$ finite. We leave the proof that $\pi \rtimes U$ is a $*$-homomorphism as Exercise 13.5.1. And, conversely, in the unital case every $*$-representation of $\mathcal{A} \cdot G$ is of the form $\pi \rtimes U$ for a covariant representation $(\pi, U)$ (Exercise 13.5.2).

Both $\mathcal{A} \rtimes_\alpha G$ and $\mathcal{A} \rtimes_\alpha^r G$ each contain a copy of $\mathcal{A}$ already in $\mathcal{A} \cdot G$, in the elements $a \cdot e$. If $\mathcal{A}$ is unital, they also contain a copy of $G$, in the elements $I_{\mathcal{A}} \cdot g$. When $\mathcal{A}$ is not unital there is no canonical copy of $G$ in the crossed product, somehow defeating the goal that the construction has. But the group will always embed as unitaries in a von Neumann algebra generated by the crossed product. Regardless, the crossed product always contains the elements $a \cdot g$.

In the trivial cases, the crossed product behaves as one would expect:

*13.5.3. Example.* When $\mathcal{A} = \mathbb{C}$ and $\alpha_g = \mathrm{id}_{\mathcal{A}}$, we get that $\mathbb{C} \rtimes_\alpha G = C^*(G)$ and $\mathbb{C} \rtimes_\alpha^r G = C^*_\lambda(G)$ (Exercise 13.5.3).

*13.5.4. Example.* When $\alpha_g = \mathrm{id}_{\mathcal{A}}$ for all $g \in G$, then $\tilde{\pi} = \pi \otimes I_{\ell^2(G)}$ and $\mathcal{A} \rtimes_\alpha G = \mathcal{A} \otimes_{\min} C^*(G)$, $\mathcal{A} \rtimes_\alpha^r G = \mathcal{A} \otimes_{\min} C^*_\lambda(G)$.

*13.5.5. Example.* When $G = \{e\}$, then $\mathcal{A} \rtimes_\alpha G = \mathcal{A}$.

Similar to Exercise 13.3.10, the reduced crossed product is a quotient of the full crossed product (Exercise 13.5.4). By definition of the norm in $\mathcal{A} \rtimes_\alpha G$, each representation $\pi \rtimes U$ is bounded, and so it extends to a representation of $\mathcal{A} \rtimes_\alpha G$.

Often times the crossed product "forgets" most information about the group. For instance if $G$ is any group and we take $\mathcal{A} = c_0(G)$ with the action $\alpha_g(f)(h) = f(g^{-1}h)$, then $c_0(G) \rtimes_\alpha G \simeq \mathcal{K}(\ell^2(G))$ (Exercise 13.5.6), so any two groups with the same cardinality produce the same crossed product algebra. Also, as $\mathcal{K}(\ell^2(G))$ is simple, it follows via Exercise 13.5.4 that $c_0(G) \rtimes_\alpha^r G = c_0(G) \rtimes_\alpha G$.

For the reduced crossed product, we see that it can be calculated via any single representation, as long as it is faithful:

**13.5.6. Proposition.**

*Let $\mathcal{A}$ be a $C^*$-algebra, $G$ a discrete group, and $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ a faithful representation. Then the norm on $\mathcal{A} \rtimes_\alpha^r G$ is given by $\|x\|_r = \|(\tilde{\pi} \rtimes (I_{\mathcal{H}} \otimes \lambda))(x)\|$.*

*Proof.* By definition (and the fact that the maps $\tilde{\beta} \rtimes \lambda$ are bounded by definition so they extend to representations of $\mathcal{A} \rtimes_\alpha^r G$) we have that the norm on $\mathcal{A} \rtimes_\alpha^r G$ is given by

$$\|x\| = \sup\{(\tilde{\beta} \rtimes \lambda)(x) : \ \beta : \mathcal{A} \to \mathcal{B}(\mathcal{H}) \text{ representation}\}.$$

We first see that it is enough to only consider faithful representations. What happens is that if $\beta$ is not faithful, we can consider the representation $\beta \oplus \gamma$, with $\gamma$ faithful, and because it is a direct sum, $\|(\tilde{\beta} \rtimes \lambda)(x)\| \leq \|((\tilde{\beta} \oplus \tilde{\gamma}) \rtimes \lambda)(x)\|$.

So fix $\beta : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ faithful (we can consider the same Hilbert space for otherwise we conjugate with an appropriate unitary). Fix $F \subset G$ finite and let $P \in \mathcal{B}(\ell^2(G))$ the orthogonal projection onto $\mathrm{span}\{\delta_g : g \in F\}$. We can see $P\mathcal{B}(\ell^2(G))P$ as $M_{|F|}(\mathbb{C})$ and so we have matrix units $\{E_{g,h}\}_{g,h \in F}$. Given

$a \in \mathcal{A}$ and $h \in F$,

$$(I_{\mathcal{H}} \otimes P)\tilde{\beta}(a)(I_{\mathcal{H}} \otimes P)(\xi \otimes \delta_h) = (I_{\mathcal{H}} \otimes P)\tilde{\beta}(a)(\xi \otimes \delta_h) = \beta(\alpha_h^{-1}(a))\xi \otimes \delta_h.$$

This allows us to think of $(I_{\mathcal{H}} \otimes P)\tilde{\beta}(a)(I_{\mathcal{H}} \otimes P)$ as

$$(I_{\mathcal{H}} \otimes P)\tilde{\beta}(a)(I_{\mathcal{H}} \otimes P) = \sum_{g \in F} \beta(\alpha_g^{-1}(a)) \otimes E_{gg}.$$

Then by Exercise 10.6.9 we can write

$$\tilde{\beta}(a) = \sum_{g \in G} \beta(\alpha_g^{-1}(a)) \otimes E_{gg},$$

with the series converging sot. We get that $\tilde{\beta}(a)$ commutes with $I_{\mathcal{H}} \otimes P$. For $a \in \mathcal{A}$ and $h \in G$ we have

$$P\lambda(h)P\delta_g = \begin{cases} \delta_{hg}, & g \in F,\ hg \in F \\ 0, & \text{otherwise} \end{cases}$$

and then, with $\tilde{P} = I_{\mathcal{H}} \otimes P$,

$$\begin{aligned} \tilde{P}\tilde{\beta}(a)(I_{\mathcal{H}} \otimes \lambda(h))\tilde{P} &= \Big(\sum_{g \in F} \beta(\alpha_g^{-1}(a)) \otimes E_{g,g}\Big)(I_{\mathcal{H}} \otimes P\lambda(h)P) \\ &= \Big(\sum_{g \in F} \beta(\alpha_g^{-1}(a)) \otimes E_{g,g}\Big)\Big(\sum_{g \in F \cap h^{-1}F} I_{\mathcal{H}} \otimes E_{hg,g}\Big) \\ &= \Big(\sum_{g \in F \cap h^{-1}F} \beta(\alpha_{h^{-1}g}^{-1}(a)) \otimes E_{hg,g}\Big) \\ &\in \beta(\mathcal{A}) \otimes M_{|F|}(\mathbb{C}). \end{aligned}$$

It follows that for $x = \sum_{g \in G} a_g \cdot g \in \mathcal{A} \cdot G$ (so finitely many $a_g$ are nonzero),

$$\begin{aligned} (I_{\mathcal{H}} \otimes P)(\tilde{\beta} \rtimes \lambda)(x)\,(I_{\mathcal{H}} \otimes P) &= \sum_{h \in G} \sum_{g \in F \cap h^{-1}F} \beta(\alpha_{h^{-1}g}^{-1}(a_h)) \otimes E_{hg,g} \\ &\in \beta(\mathcal{A}) \otimes M_{|F|}(\mathbb{C}). \end{aligned}$$

The exact same computation can be applied to $\tilde{\pi} \rtimes \lambda$. The map $\beta \otimes \mathrm{id}_{M_{|F|}(\mathbb{C})}$ is faithful by Proposition 13.4.8 (recall the $\mathcal{A} \otimes M_{|F|}$ is a C*-algebra—hence with a unique norm—by Remark 13.4.12) So $(\beta \circ \pi^{-1}) \otimes \mathrm{id}_{M_{|F|}(\mathbb{C})} : \pi(\mathcal{A}) \otimes M_{|F|}(\mathbb{C}) \to \beta(\mathcal{A}) \otimes M_{|F|}(\mathbb{C})$ is isometric. Therefore

$$\|(I_{\mathcal{H}} \otimes P)(\tilde{\beta} \rtimes \lambda)(x)\,(I_{\mathcal{H}} \otimes P)\| = \|(I_{\mathcal{H}} \otimes P)(\tilde{\pi} \rtimes \lambda)(x)\,(I_{\mathcal{H}} \otimes P)\|$$

for all $x \in \mathcal{A} \cdot G$. Then Exercise 10.6.9 gives us that

$$\|(\tilde{\beta} \rtimes \lambda)(x)\| = \|(\tilde{\pi} \rtimes \lambda)(x)\|, \qquad x \in \mathcal{A} \cdot G. \qquad \square$$

**13.5.7. Theorem. (Fell's Absorption Principle)**

*Let $G$ be a discrete group and $\gamma : G \to \mathcal{B}(\mathcal{H})$ a unitary representation. Then $\lambda \otimes \gamma$ is unitarily equivalent to $\lambda \otimes I_{\mathcal{H}}$. If $(\mathcal{A}, G, \alpha)$ is a $C^*$-dynamical system and $(\pi, U)$ is a covariant representation, then the representation $(\pi \otimes I_{\ell^2(G)}, U \otimes \lambda)$ is unitarily equivalent to the regular representation $(\tilde{\pi}, I_{\mathcal{H}} \otimes \lambda)$.*

*Proof.* Let $V : \ell^2(G) \mathbin{\bar{\otimes}} \mathcal{H} \to \ell^2(G) \mathbin{\bar{\otimes}} \mathcal{H}$ be the operator

$$V : \sum_{g \in G} \delta_g \otimes \xi_g \longmapsto \sum_{g \in G} \delta_g \otimes \gamma(g)\xi_g$$

(every element in $\ell^2(G) \mathbin{\bar{\otimes}} \mathcal{H}$ is of the form above by Lemma 13.4.3). We have

$$\Big\| \sum_{g \in G} \delta_g \otimes \gamma(g)\xi_g \Big\|^2 = \sum_{g \in G} \|\gamma(g)\xi_g\|^2 = \sum_{g \in G} \|\xi_g\|^2 = \Big\| \sum_{g \in G} \delta_g \otimes \xi_g \Big\|^2,$$

so $V$ is isometric. Linearity and surjectivity are straightforward, hence $V$ is a unitary. And we have

$$\begin{aligned} V(\lambda \otimes I_{\mathcal{H}})(g)(\delta_h \otimes \xi) &= V(\lambda(g) \otimes I_{\mathcal{H}})(\delta_h \otimes \xi) = V(\delta_{gh} \otimes \xi) = \delta_{gh} \otimes \gamma(gh)\xi \\ &= \big(\lambda(g) \otimes \gamma(g)\big)(\delta_h \otimes \gamma(h)\xi) = \big(\lambda(g) \otimes \gamma(g)\big)\, V(\delta_h \otimes \xi). \end{aligned}$$

The linearity and continuity of $V$, together with the fact that it is a unitary, give

$$V(\lambda \otimes I_{\mathcal{H}})V^* = \lambda \otimes \gamma.$$

In the crossed product case the situation is similar. We define $W : \mathcal{H} \mathbin{\bar{\otimes}} \ell^2(G) \to \mathcal{H} \mathbin{\bar{\otimes}} \ell^2(G)$ by

$$W : \sum_{g \in G} \xi_g \otimes \delta_g \longmapsto \sum_{g \in G} U_g \xi_g \otimes \delta_g.$$

This gives

$$\Big\| \sum_{g \in G} U_g \xi_g \otimes \delta_g \Big\|^2 = \sum_{g \in G} \|U_g \xi_g\|^2 = \sum_{g \in G} \|\xi_g\|^2 = \Big\| \sum_{g \in G} \xi_g \otimes \delta_g \Big\|^2$$

and hence $W$ is isometric. As above, linearity and surjectivity are straightforward and $W$ is a unitary. We have, using that $(\pi, U)$ is covariant,

$$\begin{aligned} W\tilde{\pi}(a)(\xi \otimes \delta_g) &= W(\pi(\alpha_g^{-1})(a))\xi \otimes \delta_g = U_g \pi(\alpha_g^{-1}(a))\xi \otimes \delta_g \\ &= \pi(a) U_g \xi \otimes \delta_g = (\pi \otimes I_{\ell^2(G)})(a)\, W(\xi \otimes \delta_g). \end{aligned}$$

Hence $W\tilde{\pi}(a)W^* = \pi(a) \otimes I_{\ell^2(G)}$. The proof that $W(I_{\mathcal{H}} \otimes \lambda(g))W^* = U_g \otimes \lambda(g)$ is similar. $\square$

**13.5.8. Corollary.**

*Let $(\mathcal{A}, G, \alpha)$ be a $C^*$-dynamical system and $(\pi, U)$ a covariant representation with $\pi$ faithful. Then $\mathcal{A} \rtimes_\alpha^r G \simeq C^*(\pi(\mathcal{A}) \otimes I_{\ell^2(G)}, (U \otimes \lambda)(G)) \subset \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \ell^2(G))$.*

*Proof.* By Proposition 13.5.6 we can obtain $\mathcal{A} \rtimes_\alpha^r G$ by completing $\mathcal{A} \cdot G$ with the norm induced by $\tilde{\pi} \rtimes \lambda$. So $\mathcal{A} \rtimes_\alpha^r G$ can be seen as $\mathrm{C}^*(\tilde{\pi}(\mathcal{A}), I_{\mathcal{H}} \otimes \lambda(G)) \subset \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \ell^2(G)))$. By Theorem 13.5.7 this is unitarily equivalent to $C^*(\pi(\mathcal{A}) \otimes I_{\ell^2(G)}, (U \otimes \lambda)(G))$. □

The elements $U_g \otimes \lambda(g)$ might not belong to $\mathcal{A} \rtimes_\alpha^r G$ if $\mathcal{A}$ is not unital, but $\pi(A) U_g \otimes \lambda(g)$ and $U_g \pi(A) \otimes \lambda(g)$ always do; so the next definition makes sense even if $\mathcal{A}$ is not unital.

**13.5.9. Definition.**

*A conditional expectation $\mathcal{E} : \mathcal{A} \rtimes_\alpha^r G \to \mathcal{A}$ is **equivariant** if, in the representation from Corollary 13.5.8, $\mathcal{E}((U_g \otimes \lambda(g)) X (U_g \otimes \lambda(g))^{-1}) = \alpha_g(\mathcal{E}(X))$, $X \in \mathcal{A} \rtimes_\alpha^r G$, $g \in G$.*

**13.5.10. Proposition.**

*Let $(\mathcal{A}, G, \alpha)$ be a $C^*$-dynamical system. The map $\mathcal{E} : \mathcal{A} \cdot G \to \mathcal{A}$ given by $\mathcal{E}\big(\sum_g a_g \cdot g\big) = a_e$ extends to a faithful equivariant conditional expectation $\mathcal{E} : \mathcal{A} \rtimes_\alpha^r G \to \mathcal{A}$.*

*Proof.* By Corollary 13.5.8 we may assume that

$$\mathcal{A} \rtimes_\alpha^r G = C^*(\pi(\mathcal{A}) \otimes I_{\ell^2(G)}, (U \otimes \lambda)(G)) \subset \mathcal{B}(\mathcal{H}) \otimes C_\lambda^*(G)$$

for some faithful covariant representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$. Let $\tau$ be the restriction of the canonical faithful tracial state on $L(G)$ to $C_\lambda^*(G)$ (we discussed its existence of page 972 and its faithfulness on Exercise 13.3.8). Then we let $\mathcal{E} = \mathrm{id}_{\mathcal{B}(\mathcal{H})} \mathbin{\overline{\otimes}} \tau$ be the slice map (page 979) corresponding to $\tau$. This is a conditional expectation, so in particular it is bounded. It is faithful by Proposition 13.4.7. We have by definition (using finite sums as we want elements of $\mathcal{A} \cdot G$)

$$\mathcal{E}\Big(\sum_g \pi(a_g) U_g \otimes \lambda(g)\Big) = \pi(a_e) \otimes I_{\ell^2(G)} \in \pi(\mathcal{A}) \otimes I_{\ell^2(G)},$$

since $\tau(\lambda(g)) = 0$ for all $g \neq e$. As $\mathcal{E}$ is continuous, its range stays within $\pi(\mathcal{A}) \otimes I_{\ell^2(G)}$.

Finally, for equivariance let $\tilde{X} = \sum_h \pi(a_h)U_h \otimes \lambda(h)$. Then

$$\begin{aligned}
\mathcal{E}\Big((U_g \otimes \lambda(g))\tilde{X}(U_{g^{-1}} \otimes \lambda(g^{-1}))\Big) &= \mathcal{E}\Big(\sum_h U_g\pi(a_h)U_hU_g^{-1} \otimes \lambda(ghg^{-1})\Big) \\
&= \mathcal{E}\Big(\sum_h \pi(\alpha_g(a_h))U_gU_hU_g^{-1} \otimes \lambda(ghg^{-1})\Big) \\
&= \mathcal{E}\Big(\sum_h \pi(\alpha_g(a_h))U_{ghg^{-1}} \otimes \lambda(ghg^{-1})\Big) \\
&= \mathcal{E}\Big(\sum_h \pi(\alpha_g(a_{g^{-1}hg}))U_h \otimes \lambda(h)\Big) \\
&= \pi(\alpha_g(a_e)) = \pi(\alpha_g(\mathcal{E}(\tilde{X}))).
\end{aligned}$$

By continuity the equality $\mathcal{E}((U_g \otimes \lambda(g))\tilde{X}(U_g \otimes \lambda(g))^{-1}) = \alpha_g(\mathcal{E}(\tilde{X}))$ holds for all $\tilde{X} \in \mathcal{A} \rtimes_\alpha^r G$. ☐

**13.5.11. Definition.**

*Let $X$ be a set and $G$ a group. We say that $G$* ***acts*** *on $X$ if there exists a map $\vartheta : G \times X \to X$ such that*

$$\vartheta(e, x) = x, \qquad \vartheta(s, \vartheta(t, x)) = \vartheta(st, x), \qquad s, t \in G,\ x \in X. \tag{13.27}$$

*This is often denoted $\vartheta(s, x) = s \cdot x$. When $G$ is a topological group and $X$ a topological space, the action is* ***continuous*** *if $\vartheta$ is continuous. When both $G$ and $X$ are locally compact Hausdorff, $(G, X)$ is a* ***locally compact transformation group****.*

Given a locally compact transformation group $(G, X)$, for each $s \in G$ the map $x \longmapsto s \cdot x$ is a homeomorphism. Then $\alpha : G \to \operatorname{Aut} C_0(X)$ given by

$$[\alpha_s(f)](x) = f(s^{-1} \cdot x) \tag{13.28}$$

makes $(C_0(X), G, \alpha)$ a C*-dynamical system (Exercise 13.5.5). We will consider crossed products of the form $C_0(X) \rtimes_\alpha G$, where the action is induced by a locally compact transformation group, in Section 14.6.

It turns out that every C*-dynamical system with $\mathcal{A}$ abelian comes from a group transformation:

***13.5.12. Proposition.***

*Let $(\mathcal{A}, G, \alpha)$ be a C*-dynamical system with $\mathcal{A}$ abelian. Then there exists a locally compact Hausdorff space $X$ with $\mathcal{A} = C_0(X)$ and an action of $G$ on $X$ such that*

$$[\alpha_s f](x) = f(s^{-1} \cdot x), \qquad s \in G,\ f \in C_0(X),\ x \in X. \tag{13.29}$$

*Proof.* By Theorem 11.2.3 there exists $X$ locally compact Hausdorff with $\mathcal{A} = C_0(X)$. Fix $s \in G$. As $\alpha_s \in \operatorname{Aut} C_0(X)$, by Theorem 7.4.8 there exists a unique homeomorphism $h_s$ of $X$ with

$$[\alpha_s f](x) = f(h_s(x)), \qquad x \in X.$$

From

$$f((h_s \circ h_t)(x)) = [\alpha_t \circ \alpha_s f](x) = \alpha_{ts} f(x) = f(h_{ts}(x))$$

and the uniqueness, $h_{ts} = h_s \circ h_t$. Also from the uniqueness, $h_e = \mathrm{id}_X$. Then

$$s \cdot x = h_s^{-1}(x)$$

defines an action of $G$ on $X$ that satisfies (13.29). ☐

### 13.5.1. Exercises.

(13.5.1) Let $(\mathcal{A}, G, \alpha)$ be a C*-dynamical system and $(\pi, U)$ a covariant representation. Show that $\pi \rtimes U$ is indeed a representation.

(13.5.2) Show that when $\mathcal{A}$ is unital every $*$-representation $\beta : \mathcal{A} \cdot G \to \mathcal{B}(\mathcal{H})$ is $\beta = \pi \rtimes U$ for a covariant representation $(\pi, U)$.

(13.5.3) Let $\mathcal{A} = \mathbb{C}$ and $\alpha_g(g) = I_{\mathcal{A}}$ for al $g$. Show that $\mathbb{C} \rtimes_\alpha G = C^*(G)$ and $\mathbb{C} \rtimes_\alpha^r G = C_\lambda^*(G)$.

(13.5.4) Show that $\mathcal{A} \rtimes_\alpha^r G$ is a quotient of $\mathcal{A} \rtimes_\alpha G$ via a surjection that extends the identity map on $\mathcal{A} \cdot G$.

(13.5.5) Show that (13.28) makes $(C_0(X), G, \alpha)$ a C*-dynamical system.

(13.5.6) Let $G$ be a group, $\mathcal{A} = c_0(G)$, and $\alpha_g(f)(h) = f(g^{-1}h)$. Show that $c_0(G) \rtimes_\alpha G \simeq \mathcal{K}(\ell^2(G))$.

(13.5.7) Let $(G, X)$ be a locally compact transformation group. Show that if $\operatorname{Aut} C_0(X)$ is considered with the pointwise-norm topology, then $\alpha$ as in (13.28) is continuous.

CHAPTER

# 14

# von Neumann Algebras

In Chapter 12 we saw central parts of the theory of von Neumann algebras. We also saw a way to construct certain non-trivial von Neumann algebras in Example 13.3.13. We continue here with the most common techniques and results of the theory.

## 14.1. Subalgebras

Given a $*$-algebra $\mathcal{M}$ and a projection $P \in \mathcal{M}$, we call $P\mathcal{M}P$ a **corner**. The reason is the analogy with the case $\mathcal{M} = M_n(\mathbb{C})$ and $P = E_{11} + \cdots + E_{kk}$, where $P\mathcal{M}P$ consists of those matrices with only the upper left $k \times k$ entries possibly nonzero. Most of the time, if $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ we consider $P\mathcal{M}P$ as a subalgebra of $\mathcal{B}(P\mathcal{H}) \simeq P\mathcal{B}(\mathcal{H})P$.

**14.1.1. Proposition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $P \in \mathcal{M}$, $Q \in \mathcal{M}'$ projections. Then $P\mathcal{M}P \subset \mathcal{B}(P\mathcal{H})$ and $\mathcal{M}Q \subset \mathcal{B}(Q\mathcal{H})$ are von Neumann algebras, and*

$$(P\mathcal{M}P)' = \mathcal{M}'P, \qquad (\mathcal{M}Q)' = Q\mathcal{M}'Q.$$

*Also $\mathcal{Z}(P\mathcal{M}P) = \mathcal{Z}(\mathcal{M})P$.*

*Proof.* If $T \in \mathcal{M}'$ and $S \in \mathcal{M}$, then $(TP)(PSP) = TPSP = PSTP = (PSP)TP$. So $P\mathcal{M}P \subset (\mathcal{M}'P)'$. Conversely, fix $T \in (\mathcal{M}'P)' \subset \mathcal{B}(P\mathcal{H})$. Let $T_p = TP \in \mathcal{B}(\mathcal{H})$. For any $S \in \mathcal{M}'$,

$$ST_p = STP = (SP)(TP) = (TP)(SP) = TPS = T_pS.$$

Therefore $T_p \in \mathcal{M}'' = \mathcal{M}$ and $T = PT_pP \in P\mathcal{M}P$. Thus $(\mathcal{M}'P)' \subset P\mathcal{M}P$ and hence $P\mathcal{M}P = (\mathcal{M}'P)'$. In particular $P\mathcal{M}P$ is a von Neumann algebra. We immediately get $(P\mathcal{M}P)' = (\mathcal{M}'P)''$. Thing is, we don't know yet that $\mathcal{M}'P$ is a von Neumann algebra! To address that, fix a unitary $U \in (P\mathcal{M}P)'$. We want to extend it to an operator in $\mathcal{B}(\mathcal{H})$. Let $\mathcal{K} = \overline{\mathcal{M}P\mathcal{H}}$. This is a subspace of $\mathcal{H}$ that is invariant both for $\mathcal{M}$ and $\mathcal{M}'$, and if $R$ is the orthogonal projection onto $\mathcal{K}$, then $R \in \mathcal{Z}(\mathcal{M})$ (Exercise 14.1.1). Given $T_1, \ldots, T_n \in \mathcal{M}$ and $\xi_1, \ldots, \xi_n \in \mathcal{H}$, let $\tilde{U} : \mathcal{H} \to \mathcal{H}$ be the linear operator that is zero on $\mathcal{K}^\perp$ and

$$\tilde{U}\Big(\sum_{j=1}^n T_jP\xi_j\Big) = \sum_{j=1}^n T_jPU\xi_j.$$

A first question is whether this is well-defined; but we will show that it is an isometry, which in particular guarantees that it is well-defined. We have, since $U \in (P\mathcal{M}P)'$,

$$\begin{aligned}\Big\|\tilde{U}\Big(\sum_{j=1}^n T_jP\xi_j\Big)\Big\|^2 &= \Big\|\sum_{j=1}^n T_jPU\xi_j\Big\|^2 = \sum_{k,j=1}^n \langle T_kPU\xi_k, T_jPU\xi_j\rangle \\ &= \sum_{k,j=1}^n \langle U^*PT_j^*T_kPU\xi_k, \xi_j\rangle = \sum_{k,j=1}^n \langle PT_j^*T_kP\xi_k, \xi_j\rangle \\ &= \Big\|\sum_{j=1}^n T_jP\xi_j\Big\|^2.\end{aligned}$$

So $\tilde{U}$ is isometric on $\mathcal{K}$ and hence, after extending to $\mathcal{K}$ (isometry on a dense subspace), $\tilde{U} \in \mathcal{B}(\mathcal{H})$. For any $S, T \in \mathcal{M}$ and $\xi \in \mathcal{H}$,

$$(\tilde{U}S)TP\xi = \tilde{U}STP\xi = STPU\xi = S(TPU\xi) = (S\tilde{U})TP\xi.$$

Thus $\tilde{U}S = S\tilde{U}$ on $\mathcal{K}$. And on $\mathcal{K}^\perp$, we have that $\tilde{U} = 0$, and $\mathcal{K}^\perp$ is invariant for $S$, so $\tilde{U}S = S\tilde{U}$ everywhere. This shows that $\tilde{U} \in \mathcal{M}'$. We also have, for $\xi \in P\mathcal{H}$,

$$\tilde{U}RP\xi = \tilde{U}P\xi = PU\xi = UP\xi = U\xi.$$

So $U = \tilde{U}P \in \mathcal{M}'P$. As any C$^*$-algebra is spanned by its unitaries (Remark 10.4.8), we have shown that $(P\mathcal{M}P)' \subset \mathcal{M}'P$. The reverse inclusion was shown in the first line of the proof, so $(P\mathcal{M}P)' = \mathcal{M}'P$.

The equality $(\mathcal{M}Q)' = Q\mathcal{M}'Q$ now follows trivially. Indeed, we know from the previous part of the proof, applied to $\mathcal{M}'$ and $Q$, that $Q\mathcal{M}'Q$ and $\mathcal{M}Q$ are von Neumann algebras, and that $(Q\mathcal{M}'Q)' = \mathcal{M}Q$. Taking commutants, $(\mathcal{M}Q)' = Q\mathcal{M}'Q$.

We will prove the equality $\mathcal{Z}(P\mathcal{M}P) = \mathcal{Z}(\mathcal{M})P$ in Corollary 14.1.4. $\square$

We know from Corollary 12.3.9 that arbitrary unions and intersections of projections in $\mathcal{M}$ stay in $\mathcal{M}$. Because of this, the following definition makes sense:

**14.1.2. Definition.**

*Let $\mathcal{M}$ be a von Neumann algebra and $P \in \mathcal{M}$ a projection. The* ***central support*** *of $P$ is the smallest projection $c(P) \in \mathcal{Z}(\mathcal{M})$ such that $P \leq c(P)$.*

For an easy example to visualize the central support, suppose that $\mathcal{M} = M_2(\mathbb{C}) \oplus M_3(\mathbb{C})$ and let $P = E_{11} \oplus 0$. Then $c(P) = I_2 \oplus 0$. If $Q = E_{22} \oplus E_{11}$, then $c(Q) = I_{\mathcal{M}}$.

We recall the notation from Definition 12.3.13.

***14.1.3. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $P \in \mathcal{M}$ a projection. Then $c(P) = [\mathcal{M}P\mathcal{H}]$.*

*Proof.* Let $Q = [\mathcal{M}P\mathcal{H}]$. We know from Exercise 14.1.1 that $Q \in \mathcal{Z}(\mathcal{M})$. We also have $P\mathcal{H} \subset \mathcal{M}P\mathcal{H}$, which shows that $P \leq Q$. Given a projection $R \in \mathcal{Z}(\mathcal{M})$ with $P \leq R$, we get $R(\mathcal{M}P\mathcal{H}) = \mathcal{M}RP\mathcal{H} \subset \mathcal{M}P\mathcal{H}$. It follows that $QR = Q$; this shows that $Q \leq R$ and hence $Q$ is the smallest central projection in $\mathcal{M}$ with $P \leq Q$. $\square$

A simple observation is that, for any von Neumann algebra,

$$\mathcal{Z}(\mathcal{M}) = \mathcal{M} \cap \mathcal{M}' = \mathcal{Z}(\mathcal{M}').$$

Next is a crucial technical result that will be used several times.

***14.1.4. Corollary.***

*Let $\mathcal{M}$ be a von Neumann algebra and $P \in \mathcal{M}$ a nonzero projection. We have a $\sigma$-weak continuous $*$-isomorphism $\gamma : \mathcal{M}'c(P) \to \mathcal{M}'P$ given by $\gamma : T \longmapsto TP$. As a consequence, $\mathcal{Z}(P\mathcal{M}P) = \mathcal{Z}(\mathcal{M})P$, and if $\mathcal{M}$ is a factor then $P\mathcal{M}P$ and $\mathcal{M}'P$ are factors in $\mathcal{B}(P\mathcal{H})$.*

*Proof.* It is straightforward that $\gamma$ is linear. Since $P \in \mathcal{M}$, it commutes with all $T \in \mathcal{M}'$, and together with the fact that $P^2 = P$, this gives that $\gamma$ is multiplicative; and from $P^* = P$ we get that $\gamma$ preserves adjoints, so it is a

$*$-homomorphism. Surjectivity is apparent. Suppose that $\gamma(T) = 0$ for some $T \in \mathcal{M}'$; that is, $TP = 0$. For any $\xi \in \mathcal{H}$ and $S \in \mathcal{M}$,

$$T(SP\xi) = S(TP)\xi = 0.$$

So $T = 0$ on $\overline{\mathcal{M}P\mathcal{H}}$, which is to say that $Tc(P) = 0$; therefore $\gamma$ is injective. It is then $\sigma$-weak continuous by Corollary 12.6.12.

Regarding the centre, it is clear that $\mathcal{Z}(\mathcal{M})c(P) \subset \mathcal{Z}(\mathcal{M}c(P))$. Fix $T \in \mathcal{Z}(\mathcal{M}\,c(P))$. This in particular implies that $T = Tc(P)$. Given $S \in \mathcal{M}$,

$$TS = Tc(P)\,S = T(Sc(P)) = Sc(P)T = ST.$$

So $T \in \mathcal{Z}(\mathcal{M})$ and hence $T \in \mathcal{Z}(\mathcal{M})c(P)$. Applying this to $\mathcal{M}'$,

$$\mathcal{Z}(\mathcal{M}'\,c(P)) = \mathcal{Z}(\mathcal{M}')c(P).$$

The isomorphism $\gamma$ then tells us (using also Proposition 14.1.1) that

$$\mathcal{Z}(P\mathcal{M}P) = \mathcal{Z}(\mathcal{M}'P) = \mathcal{Z}(\gamma(\mathcal{M}'c(P)) = \gamma(\mathcal{Z}(\mathcal{M}'c(P)))$$

$$= \gamma(\mathcal{Z}(\mathcal{M}')c(P)) = \mathcal{Z}(\mathcal{M}')P.$$

(to help with the information overload, note that the last equality is trivial given how $\gamma$ is defined).

When $\mathcal{M}$ is a factor, so is $\mathcal{M}'$ (because $\mathcal{Z}(\mathcal{M}') = \mathcal{Z}(\mathcal{M})$) and therefore $\mathcal{M}'P$ is a factor, since $c(P) = I_{\mathcal{M}}$ and $\mathcal{M}'P \simeq \mathcal{M}'c(P) = \mathcal{M}'$. Now Proposition 14.1.1 tells us that $(P\mathcal{M}P)'$ is a factor, and hence $P\mathcal{M}P$ is a factor. $\square$

**14.1.1. Exercises.**

(14.1.1) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $P \in \mathcal{M}$ a projection, $\mathcal{K} = \overline{\mathcal{M}P\mathcal{H}}$, and $Q \in \mathcal{B}(\mathcal{H})$ the orthogonal projection onto $\mathcal{K}$. Show that $\mathcal{K}$ is invariant for $\mathcal{M}$ and $\mathcal{M}'$, and conclude that $Q \in \mathcal{Z}(\mathcal{M})$.

(14.1.2) Let $\mathcal{M}$ be a factor, $S \in \mathcal{M}$, $T \in \mathcal{M}'$. Show that $ST = 0$ if and only if $S = 0$ or $T = 0$.

(14.1.3) Let $P, Q, Z \in \mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be projections with $Z$ central and $Q = ZP$. Show that $c(Q) = Zc(P)$.

(14.1.4) Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{R} \subset \mathcal{M}$ a subset that is closed under multiplication and taking adjoints, and such that $W^*(\mathcal{R}) = \mathcal{M}$. Let $P \in \mathcal{M} \cup \mathcal{M}'$ be a projection. Show that $W^*(P\mathcal{R}P) = P\mathcal{M}P$. Show also that the result is not necessarily true if $\mathcal{R}$ is not closed under multiplication.

(14.1.5) Let $\mathcal{M}$ be a von Neumann algebra, $\mathcal{H}$ a Hilbert space, $P \in \mathcal{M}$ a projection. Fix also an orthonormal basis for $\mathcal{H}$ and let $\{E_{kj}\}$ be the associated matrix units. Show that

$$P\mathcal{M}P \simeq (P \otimes E_{11})(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H}))(P \otimes E_{11}).$$

## 14.2. Comparison of Projections

A main tool in working with von Neumann algebras is the study of its projections. We already know from Corollary 12.4.16 that any von Neumann algebra is the closed linear span of its projections, so lots of them are at hand. When a von Neumann algebra $\mathcal{M}$ is finite-dimensional, the analysis we did in Section 11.8 applies; and it was heavily based on studying the projections, and we also used partial isometries for that. In hindsight the use of partial isometries is not a surprise, because a partial isometry is an algebraic tool to say that two subspaces have the same size. This will hopefully make the coming definitions appear more natural.

We defined partial isometries in Definition 10.4.9 and saw their most basic properties in Proposition 10.4.10. We know from Corollary 12.3.8 that any von Neumann algebra has lots of partial isometries. Projections are partial isometries. Unitaries are partial isometries. And in any non-abelian von Neumann algebra, partial isometries that are neither projections nor unitaries exist.

For our manipulations with projections, it is useful to keep in mind that if a sum of projections is a projection, then the projections are pairwise orthogonal (Proposition 10.5.5).

**14.2.1. Definition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Two projections $P, Q \in \mathcal{M}$ are* ***equivalent*** *if there exists a partial isometry $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = Q$. We say that $P \preceq Q$ if there exists a partial isometry $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* \leq Q$.*

*We write $P \prec Q$ to mean that $P \preceq Q$ but $P$ and $Q$ are not equivalent.*

As a couple of observations we mention that we do not have to require $V$ to be a partial isometry in Definition 14.2.1; it follows automatically from the fact that $P$ (or $Q$) is a projection (via Proposition 10.4.10). Also, a crucial fact is the requirement that $V \in \mathcal{M}$. For instance if $\mathcal{M} \subset M_2(\mathbb{C})$ is the diagonal algebra, then $P = E_{11}$ and $E_{22}$ are not equivalent, for the partial isometry that would work ($V = E_{21}$) is not in $\mathcal{M}$. The same two projections are equivalent in $\mathcal{M} = M_2(\mathbb{C})$.

When $\mathcal{M}$ is finite-dimensional, equivalence of projections can be easily characterized (Exercises 14.2.2 and 14.2.3). We mention, as a warm-up exercise, that $P \preceq 0$ if and only if $P = 0$ (Exercise 14.2.1).

*14.2.2. Remark*. A byproduct of Proposition 10.4.10 that we will use often is that if $V^*V = P$ and $VV^* = Q$, then $V = QVP$. A meaningful consequence of this is that the partial isometry that makes $P$ and $Q$ equivalent always lives in $(P \vee Q)\mathcal{M}(P \vee Q)$.

**14.2.3. Proposition.**

*Let $\mathcal{M}$ be a von Neumann algebra. Equivalence of projections is an equivalence relation.*

*Proof.* It is entirely straightforward that the relation is reflexive and symmetric. What is not equally obvious is that it is transitive. Suppose that $P \sim Q$ and $Q \sim R$. Then there exist partial isometries $V, W \in \mathcal{M}$ with $V^*V = P$, $VV^* = Q$, $W^*W = Q$, $WW^* = R$. Let $U = WV \in \mathcal{M}$. Then

$$U^*U = V^*W^*WV = V^*QV = V^*V = P,$$

$$UU^* = WVV^*W^* = WQW^* = WW^* = R.$$

So $P \sim R$ and the relation is transitive. □

It is not true in general that a product of partial isometries is a partial isometry. But in the case of the proof of Proposition 14.2.3 the initial and final spaces align nicely to make the product work.

We want to use $\preceq$ as an order in the set of projections of a von Neumann algebra $\mathcal{M}$. It is obviously reflexive, and transitivity is not hard (Exercise 14.2.4). But it is not antisymmetric: if $P \sim Q$ then $P \preceq Q$ and $Q \preceq P$ even when $P \neq Q$. Still, the next best thing holds; that is, if $P \preceq Q$ and $Q \preceq P$, then $P \sim Q$. The proof of this is not straightforward, but rather a tedious almost line-by-line rehash of the idea in the proof of Schröder–Bernstein (Exercise 14.2.6).

Let us revisit a result.

**14.2.4. Proposition. (Kaplansky's Formula in a von Neumann Algebra)**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, and $P, Q \in \mathcal{M}$ projections. Then*

$$P \vee Q - Q \sim P - P \wedge Q. \tag{14.1}$$

*Proof.* The proof of Proposition 10.5.11 applies, and we get that $V \in \mathcal{M}$ via Corollary 12.3.8. □

We defined the central support of a projection $P \in \mathcal{M}$ in Definition 14.1.2.

***14.2.5. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra and $P, Q \in \mathcal{M}$ projections with $P \preceq Q$. Then $c(P) \leq c(Q)$. In particular, if $P \sim Q$ then $c(P) = c(Q)$.*

*Proof.* By hypothesis there exists a partial isometry $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* \leq Q$. Using Proposition 14.1.3,

$$\mathcal{M}P\mathcal{H} = \mathcal{M}V^*V\mathcal{H} \subset \mathcal{M}V^*\mathcal{H} = \mathcal{M}V^*Q\mathcal{H} \subset \mathcal{M}Q\mathcal{H}.$$

So $\overline{\mathcal{M}P\mathcal{H}} \subset \overline{\mathcal{M}Q\mathcal{H}}$ and thus $c(P) \leq c(Q)$. When $P \sim Q$ we get both $c(P) \leq c(Q)$ and $c(Q) \leq c(P)$, so $c(P) = c(Q)$. □

Proposition 14.2.5 shows a reason why, as mentioned above, $E_{11}$ and $E_{22}$ are not equivalent in $\mathbb{C} \oplus \mathbb{C} \subset M_2(\mathbb{C})$. They are their own central supports, and they are orthogonal to each other. On the other hand, when considered in the algebra $M_2(\mathbb{C})$, both have central support $I_2$.

The next result is technical, but crucial. We used a finite-dimensional version of it in Lemma 11.8.5 and Theorem 11.8.10.

***14.2.6. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra and $P, Q \in \mathcal{M}$ projections. The following statements are equivalent:*

*(i) there exist nonzero projections $P_0, Q_0 \in \mathcal{M}$ with $P_0 \leq P$, $Q_0 \leq Q$, and $P_0 \sim Q_0$;*

*(ii) $c(P)c(Q) \neq 0$;*

*(iii) there exists $T \in \mathcal{M}$ with $QTP \neq 0$.*

*Proof.* (i) $\implies$ (ii) Suppose that $c(P)c(Q) = 0$. If $P_0 \sim Q_0$ with $P_0 \leq P$ and $Q_0 \leq Q$, from Proposition 14.2.5 we have $c(P_0) = c(Q_0)$. We also have $c(P_0) \leq c(P)$ and $c(Q_0) \leq c(Q)$; so $c(P_0) = c(P_0)c(P)c(Q) = 0$, and similarly $c(Q_0) = 0$. Therefore $P_0 = Q_0 = 0$.

(ii) $\implies$ (iii) As $c(P)c(Q) = c(P) \wedge c(Q)$ (because they commute) we have that $\overline{\mathcal{M}P\mathcal{H}} \cap \overline{\mathcal{M}Q\mathcal{H}} \neq \{0\}$. Given nonzero $\xi$ in the intersection, there exist sequences $\{R_nP\nu_n\}$ and $\{S_nQ\eta_n\}$ with $R_nP\nu_n \to \xi$ and $S_nQ\eta_n \to \xi$. So

$$0 < \|\xi\|^2 = \lim_n \langle R_nP\nu_n, S_nQ\eta_n\rangle.$$

This allows us to choose $R, S \in \mathcal{M}$ and $\nu, \eta \in \mathcal{H}$ with $\langle RP\nu, SQ\eta\rangle > 0$. We can write this as $\langle QS^*RP\nu, \eta\rangle > 0$. Taking $T = S^*R \in \mathcal{M}$, we have $QTP \neq 0$.

(iii) $\Longrightarrow$ (i) As $QTP \in \mathcal{M}$, if $QTP = V|QTP|$ is its polar decomposition, then $V \in \mathcal{M}$ by Corollary 12.3.8. By Proposition 10.4.11, $P_0 = V^*V \in \mathcal{M}$ is the range projection of $(QTP)^*$ and $Q_0 = VV^*$ is the range projection of $QTP$. So $P_0 \leq P$, $Q_0 \leq Q$, nonzero, and $P_0 \sim Q_0$. ☐

**14.2.7. Proposition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\{P_j\}, \{Q_j\} \subset \mathcal{M}$ two families of pairwise orthogonal projections in $\mathcal{M}$, such that $P_j \sim Q_j$ for all $j$. Then $\sum_j P_j \sim \sum_j Q_j$.*

*Proof.* By hypothesis, for each $j$ there exists $V_j \in \mathcal{M}$ with $V_j^* V_j = P_j$ and $V_j V_j^* = Q_j$. We define $V : \mathcal{H} \to \mathcal{H}$ by $V|_{P_j\mathcal{H}} = V_j$ for all $j$, and $V = 0$ on $\left(\bigcup_j P_j\mathcal{H}\right)^\perp$. Then $V^*V = \sum_j P_j$ and $VV^* = \sum_j Q_j$. What remains to check is that $V \in \mathcal{M}$. Let $T \in \mathcal{M}'$. Fix $j$ and $\xi \in P_j\mathcal{H}$. Then

$$\begin{aligned} TV\xi &= TVP_j\xi = TV_j\xi = V_jT\xi = V_jTP_j\xi \\ &= V_jP_jT\xi = VP_jT\xi = VTP_j\xi = VT\xi. \end{aligned}$$

As this can be done for any $j$ and any $\xi \in P_j\mathcal{H}$, we have that $VT = TV$ on $\mathcal{K} = \overline{\operatorname{span}} \bigcup_j P_j\mathcal{H}$. Given $\eta \in \mathcal{K}^\perp$ and $\xi \in P_j\mathcal{H}$,

$$\langle T\eta, \xi\rangle = \langle \eta, T^*\xi\rangle = \langle \eta, T^*P_j\xi\rangle = \langle \eta, P_jT^*\xi\rangle = 0.$$

So $T$ leaves $\mathcal{K}^\perp$ invariant. For any $\eta \in \mathcal{K}^\perp$, $T\eta \in \mathcal{K}^\perp$ so $VT\eta = 0$. And $TV\eta = 0$ too. So $VT = TV$ on all of $\mathcal{H}$. Hence $V \in \mathcal{M}'' = \mathcal{M}$. ☐

We have seen that there are examples of non-comparable projections. But we have the next best thing:

**14.2.8. Theorem. (Comparison)**

*Let $\mathcal{M}$ be a von Neumann algebra and $P, Q \in \mathcal{M}$ projections. Then there exists a projection $Z \in \mathcal{Z}(\mathcal{M})$ such that $ZP \preceq ZQ$ and $(I_\mathcal{M} - Z)Q \preceq (I_\mathcal{M} - Z)P$.*

*Proof.* We use Zorn's Lemma to get families $\{P_j\}_{j\in J}$, $\{Q_j\}_{j\in J}$ of pairwise orthogonal projections in $\mathcal{M}$, maximal with respect to $P_j \leq P$, $Q_j \leq Q$, $P_j \sim Q_j$ for all $j \in J$. Then Proposition 14.2.7 gives us that

$$P_0 = \sum_j P_j \sim \sum_j Q_j = Q_0$$

in $\mathcal{M}$. Let $Z = c(Q - Q_0)$. We claim that $ZP = ZP_0$. If this is the case, we have $ZP = ZP_0 \sim ZQ_0 \leq ZQ$. And

$$(I_{\mathcal{M}} - Z)Q = (I_{\mathcal{M}} - Z)Q_0 + (I_{\mathcal{M}} - Z)(Q - Q_0) = (I_{\mathcal{M}} - Z)Q_0$$
$$\sim (I_{\mathcal{M}} - Z)P_0 \leq (I_{\mathcal{M}} - Z)P.$$

So it remains to prove that $ZP = ZP_0$. Let $P_1 \leq ZP - ZP_0$ be a projection. Then $c(P_1) \leq Z = c(Q - Q_0) \leq c(Q)$. So $c(P_1)c(Q) = c(P_1)$. If $P_1 \neq 0$, then by Proposition 14.2.6 there exists $Q_1 \leq Q$ with $P_1 \sim Q_1$ and $Q_1 \leq c(Q - Q_0)$; but this contradicts the maximality of the families. Therefore $ZP = ZP_0$. $\square$

There is a more granular version of the Comparison Theorem where one gets central projections $Z_1, Z_2, Z_3$ with $Z_1 + Z_2 + Z_3 = I_{\mathcal{M}}$ and $Z_1 P \prec Z_1 Q$, $Z_2 P \sim Z_2 Q$, and $Z_3 Q \prec Z_3 P$. In such case the three projections are unique.

One of the reasons factors are important is precisely comparison:

**14.2.9. Corollary.**

*Let $\mathcal{M}$ be a factor, and $P, Q \in \mathcal{M}$ projections. Then either $P \sim Q$, $P \prec Q$, or $Q \prec P$.*

*Proof.* By Theorem 14.2.8, there exists $Z \in \mathcal{Z}(\mathcal{M}) = \mathbb{C}\, I_{\mathcal{M}}$ with $ZP \preceq ZQ$ and $(I_{\mathcal{M}} - Z)Q \preceq (I_{\mathcal{M}} - Z)P$. As $Z$ is a scalar, it is either 0 or $I_{\mathcal{M}}$, which tells us that we have at least one of $P \preceq Q$ and $Q \preceq P$. $\square$

So, in factor, any two projections are comparable. We have already seen this in the finite-dimensional case (Exercise 14.2.2).

Here is another nice characterization of the central support:

**14.2.10. Proposition.**

*Let $\mathcal{M}$ be a von Neumann algebra and $E \in \mathcal{M}$ a projection. Let $\{P_j\}$ a family of pairwise orthogonal projections, maximal with respect to the property $P_j \preceq E$. Then $c(E) = \sum_j P_j$.*

*Proof.* Let $Z = \sum_j P_j$. By Comparison there exists $Q$ central with

$$Q(I_{\mathcal{M}} - Z) \preceq QE.$$

If $Q(I_{\mathcal{M}} - Z) \neq 0$ we contradict the maximally of $\{P_j\}$. So $Q \leq Z$. We also have from Comparison that

$$(I_{\mathcal{M}} - Q)E \preceq (I_{\mathcal{M}} - Q)(I_{\mathcal{M}} - Z) = I_{\mathcal{M}} - Z.$$

This forces $(I_{\mathcal{M}} - Q)E = 0$, for otherwise we again contradict the maximality. Thus $E \leq Q$. As $Q$ is central, we have $c(E) \leq Q$. For each $j$ we have $P_j \preceq E$, so by Proposition 14.2.5 $P_j \leq c(E) \leq Q$. This implies that $Z \leq Q$, and we conclude that $Z = Q$.

At this stage we have that $Z$ is central and $E \leq c(E) \leq Z$. The projection $Z' = Z - c(E)$ is central and $Z'E = 0$. Therefore, for any $j$ we have $Z'P_j \preceq Z'E = 0$ (Exercise 14.2.5), so $Z'P_j = 0$. We conclude that $Z' = Z'Z = 0$, and hence $Z = c(E)$. $\square$

It is worth mentioning that Proposition 14.2.10 does not say that the projections $\{P_j\}$ are pairwise equivalent. For instance if $\mathcal{M} = M_3(\mathbb{C})$ and $E = E_{11} + E_{22}$, then a maximal family would be $\{E, E_{33}\}$.

Certain characteristics of projections will help us understand von Neumann algebras better.

**14.2.11. Definition.**

*Let $\mathcal{M}$ be a von Neumann algebra, and $P \in \mathcal{M}$ a projection. We say that $P$ is*

- ***minimal*** *if $P\mathcal{M}P = \mathbb{C}\, I_{\mathcal{M}}$;*
- ***abelian*** *if the algebra $P\mathcal{M}P$ is abelian;*
- ***infinite*** *if there exists $Q \leq P$ with $Q \neq P$ and $Q \sim P$*
- ***finite*** *if it is not infinite;*
- ***semifinite*** *if there exists a family $\{P_j\} \subset \mathcal{M}$ of pairwise orthogonal finite projections with $P = \sum_j P_j$;*
- ***purely infinite*** *if $Q \leq P$ for $Q \in \mathcal{M}$ implies that $Q$ is infinite;*
- ***properly infinite*** *if $ZP$ is infinite for all nonzero projections $Z \in \mathcal{Z}(\mathcal{M})$;*
- *the algebra $\mathcal{M}$ is **finite**, **semifinite**, **infinite**, **purely infinite**, or **properly infinite** if $I_{\mathcal{M}}$ is;*
- *the algebra $\mathcal{M}$ is **diffuse** if every projection in $\mathcal{M}$ admits a proper subprojection.*

Examples of most of these situations are easy to come by. For instance take $\mathcal{M} = M_2(\mathbb{C})$, and $P$ any rank-one projection; then $P$ is minimal, abelian, and finite (Exercise 14.2.13). If we take $\mathcal{M} = \mathcal{B}(\mathcal{H})$ for infinite-dimensional $\mathcal{H}$, then $I_{\mathcal{H}}$ is infinite, for $I_{\mathcal{H}} \sim I_{\mathcal{H}} - E_{11}$: we have $I_{\mathcal{H}} = V^*V$, $I_{\mathcal{H}} - E_{11} = VV^*$ with $V$ the unilateral shift (Exercise 14.2.22). So $\mathcal{B}(\mathcal{H})$ is infinite, and the minimal projections are precisely the rank-one projections, which are also the abelian projections.

At this stage we do not have examples of von Neumann algebras which lack finite projections, so we cannot give examples of purely infinite projections; in terms to be defined soon, purely infinite projections only exist when $\mathcal{M}$ has a type III central block. What is clear is that a purely infinite projection is properly infinite. As for properly infinite projections, in a factor these are just the infinite projections, so an example of a properly infinite projection is $I_{\mathcal{H}} \in \mathcal{B}(\mathcal{H})$ when $\dim \mathcal{H} = \infty$. For an example of an infinite projection which is not properly infinite we need non-trivial centre. Consider $\mathcal{M} = \mathbb{C} \oplus \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$. Then $I_{\mathcal{M}}$ is infinite, but it is not properly infinite since $(1 \oplus 0)I_{\mathcal{M}} = 1 \oplus 0$, which is finite.

It is important to emphasize again that equivalence of projections is determined relative to the algebra. For example let $\mathcal{M} = L^\infty[0,1]$. In this algebra, any projection is infinitely divisible, since it will always be the characteristic function of an infinite set. And $\mathcal{M}$ can only be represented faithfully on infinite-dimensional Hilbert spaces. But $\mathcal{M}$ is finite! Indeed, since it is abelian, no two distinct projections are equivalent: if $P = V^*V$ and $Q = VV^*$, then $Q = VV^* = V^*V = P$. So every projection is finite, and in particular $\mathcal{M}$ is finite. With the same argument, every abelian von Neumann algebra is finite, and every abelian projection is finite. Non-commutative infinite-dimensional finite von Neumann algebras exist, but we will need more machinery to discuss them; we do have one example at hand, though: that would be the one from Example 13.3.13. We know from Exercise 13.3.8 that the tracial state discussed there is faithful. If $P, Q \in L(G)$ with $G$ icc and $Q \leq P$, $Q \sim P$, there exists $V \in L(G)$ with $V^*V = P$, $VV^* = Q$; then $\tau(P - Q) = \tau(V^*V) - \tau(VV^*) = 0$ and, as $\tau$ is faithful, we get that $Q = P$. So every projection in $L(G)$ is finite.

It is worth mentioning that if $P \sim Q$ in $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, then $P \sim Q$ in $\mathcal{B}(\mathcal{H})$ (because the partial isometry is in $\mathcal{M}$ and hence in $\mathcal{B}(\mathcal{H})$), but the converse is not necessarily true, as seen above.

For an example where abelian projections are not minimal, consider $\mathcal{M} = M_2(L^\infty[0,1])$. The projection $E_{11} \in \mathcal{M}$ is abelian but not minimal.

As one would expect, if $P \in \mathcal{M}$ is a projection satisfying any of the properties in Definition 14.2.11 and $Q \in \mathcal{M}$ with $Q \sim P$, then $Q$ satisfies the same property (Exercise 14.2.16).

Since equivalence of projections can only happen within the same central block (in the sense that $P \sim Q$ implies $c(P) = c(Q)$, Proposition 14.2.5), it is possible that a sum of infinitely many projections is finite:

***14.2.12. Lemma.***

*Let $\{P_j\} \subset \mathcal{M}$ be a family of projections such that $\{c(P_j)\}$ are pairwise orthogonal, and let $P = \sum_j P_j$. If every $P_j$ is abelian/finite, then $P$ is abelian/finite.*

*Proof.* Assume first that each $P_j$ is abelian. If $k \neq j$, from Proposition 14.2.6 we know that $P_kTP_j = 0$ for all $T \in \mathcal{M}$. Then

$$PTP = \sum_{k,j} P_kTP_j = \sum_j P_jTP_j.$$

Now

$$\begin{aligned}(PTP)(PSP) &= \sum_{k,j} P_kTP_kP_jSP_j = \sum_j P_jTP_jSP_j \\ &= \sum_j P_jSP_jTP_j = (PSP)(PTP).\end{aligned}$$

So $P\mathcal{M}P$ is abelian.

If every $P_j$ is finite, suppose that $Q \leq P$ and $Q \sim P$. So $V^*V = P$ and $VV^* \leq P$. Let $V_j = c(P_j)V$. Because the central supports are pairwise orthogonal, $V_j^*V_j = c(P_j)V^*V = c(P_j)P = P_j$. And $V_jV_j^* = c(P_j)VV^* \leq c(P_j)P = P_j$. As $P_j$ is finite, we get that $V_jV_j^* = P_j$. Then

$$Q = VV^* = \sum_j c(P_j)V_jV_j^* = \sum_j P_j = P,$$

and $P$ is finite. □

A consequence of Lemma 14.2.12 is that every projection can written as the sum of two projections, with orthogonal central supports, one finite and the other properly infinite.

**14.2.13. Proposition.**

*Let $\mathcal{M}$ be a von Neumann algebra, $P \in \mathcal{M}$ a projection. Then there exists a central projection $Z \in \mathcal{Z}(\mathcal{M})$ with $ZP$ finite and $(I_{\mathcal{M}} - Z)P$ properly infinite. This decomposition is unique.*

*Proof.* If $P$ is properly infinite, we take $Z = 0$. Otherwise, via Zorn's Lemma we can construct a maximal pairwise orthogonal family $\{Z_j\}$ of central projections such that $Z_jP$ is finite for all $j$. Let $Z = \sum_j Z_j$. By Lemma 14.2.12, $ZP$ is finite. And $(I_{\mathcal{M}} - Z)P$ is properly infinite, for if $Z_0$ is a nonzero central projection with $Z_0(I_{\mathcal{M}} - Z)P$ finite, this contradicts the maximality of the family $\{Z_j\}$

As for the uniqueness, suppose that $Z' \in \mathcal{Z}(\mathcal{M})$ is a projection with $Z'P$ finite and $(I_{\mathcal{M}} - Z')P$ properly infinite. Then $Z'(I_{\mathcal{M}} - Z)P$, if nonzero, would be both finite and properly infinite. So it has to be that $Z'(I_{\mathcal{M}} - Z)P = 0$. Then $Z'P = Z'ZP + Z'(I_{\mathcal{M}} - Z)P = Z'ZP = ZP$ (the last equality by exchanging roles). □

Halving is likely the most useful tool when dealing with infinite projections.

**14.2.14. Proposition.** *(Halving)*

*Let $P \in \mathcal{M}$ be a properly infinite projection. Then there exists a projection $Q \in \mathcal{M}$ such that $Q \leq P$ and $Q \sim P \sim P - Q$.*

*Proof.* By working on $P\mathcal{M}P$, we may assume that $P = I_{\mathcal{M}}$. When $I_{\mathcal{M}}$ is properly infinite, so is every central projection by definition. We claim—to be proven—that for any nonzero projection $Z \in \mathcal{Z}(\mathcal{M})$ there exists a nonzero projection $R \in \mathcal{M}$ with $R \leq Z$, and $R \sim c(R) - R \sim c(R)$. This allows us, via Zorn's Lemma, to have a family of projections $\{R_j\} \subset \mathcal{M}$, maximal with respect to the property that $\{c(R_j)\}$ are pairwise orthogonal, and $R_j \sim c(R_j) - R_j \sim c(R_j)$. If we put $Q = \sum_j R_j$, then $c(Q) = \sum_j c(R_j)$. If $c(Q) \neq I_{\mathcal{M}}$, by the claim (applied to $I_{\mathcal{M}} - c(Q)$) we would contradict the maximality. So $c(Q) = I_{\mathcal{M}}$, and using Proposition 14.2.7

$$Q = \sum_j R_j \sim \sum_j c(R_j) - R_j = I_{\mathcal{M}} - Q \sim \sum_j c(R_j) = I_{\mathcal{M}}.$$

So it remains to prove the claim. Let $Z \in \mathcal{Z}(\mathcal{M})$, nonzero. As it is not finite (because $I_{\mathcal{M}}$ is properly infinite), there exists $V \in \mathcal{M}$ with $V^*V = Z$ and $VV^* \leq Z$, $VV^* \neq Z$. Define recursively $P_0 = Z - VV^*$ and $P_{n+1} = VP_nV^*$. Then $P_{n+1} \sim P_n$, since $P_nV^*VP_n = P_nZP_n = P_n$, and $VP_nV^* = P_{n+1}$. Also,

$$\begin{aligned} P_{n+k}P_n &= V^kP_nV^{*k}P_n = V^{n+k}P_0V^{*(n+k)}V^nP_0V^{*n} \\ &= V^{n+k}P_0V^{*k}P_0V^{*n} = 0 \end{aligned}$$

(since $V^*P_0 = 0$). So the projections $\{P_n\}$ are pairwise orthogonal and equivalent. Using Zorn's Lemma we extend $\{P_n\}$ to a maximal family $\{Q_j\}_{j\in J} \subset \mathcal{M}$ of pairwise orthogonal and pairwise equivalent projections such that $Q_j \leq Z$ for all $j$. The index set $J$ is infinite since $J \supset \mathbb{N}$. Using Proposition 1.6.29, let $J_0 \subset J$ with $|J_0| = |J| = |J \setminus J_0|$. From Proposition 14.2.7

$$\sum_{j\in J} Q_j \sim \sum_{j\in J_0} Q_j \sim \sum_{j\in J\setminus J_0} Q_j \sim \sum_{j\in J\setminus\{j_0\}} Q_j \tag{14.2}$$

for some fixed $j_0 \in J$. At this stage it is tempting to think we are done, for (14.2) looks an awful lot like $P \sim Q \sim P - Q$. The problem is that we cannot guarantee that the series adds to a prescribed projection. At this stage of the proof, we are trying to prove the claim that there exists $R \leq Z$ with $R \sim c(R) - R \sim c(R)$. We set $Q_0 = Z - \sum_{j\in J} Q_j$. By Comparison there exists a projection $Z_0 \in \mathcal{Z}(\mathcal{M})$ with $Z_0Q_0 \preceq Z_0Q_{j_0}$ and $(I_{\mathcal{M}} - Z_0)Q_{j_0} \preceq (I_{\mathcal{M}} - Z_0)Q_0$. If we had $Z_0 = 0$ we would have a subprojection $Q_0' \leq Q_0$

with $Q_0 \sim Q_{j_0}$; but this contradicts the maximality of $\{Q_j\}$. Knowing that $Z_0 \neq 0$,

$$\begin{aligned} ZZ_0 &= Q_0 Z_0 + \sum_{j \in J} Q_j Z_0 \sim Q_0 Z_0 + \sum_{j \in J \setminus \{j_0\}} Q_j Z_0 \\ &\preceq Q_{j_0} Z_0 + \sum_{j \in J \setminus \{j_0\}} Q_j Z_0 = \sum_{j \in J} Q_j Z_0 \\ &\leq ZZ_0. \end{aligned}$$

By Exercise 14.2.6, $ZZ_0 \sim \sum_{j \in J} Q_j Z_0$. Let $R = \sum_{j \in J_0} Q_j Z_0$. We have $R \sim ZZ_0$ by (14.2), and

$$\begin{aligned} ZZ_0 - R &= ZZ_0 - \sum_{j \in J} Q_j Z_0 + \sum_{j \in J \setminus J_0} Q_j Z_0 \\ &= Q_0 Z_0 + \sum_{j \in J \setminus J_0} Q_j Z_0 \\ &\sim Q_0 Z_0 + \sum_{j \in J} Q_j Z_0 = ZZ_0. \end{aligned}$$

We finish by noticing that $ZZ_0 = c(R)$, since $ZZ_0 \sim R$ and $ZZ_0$ is central (Proposition 14.2.5). □

An easy example of an infinite projection that cannot be halved is the one mentioned earlier, $1 \oplus I_{\mathcal{H}} \in \mathbb{C} \oplus \mathcal{B}(\mathcal{H})$ (Exercise 14.2.23).

Curiously, our new knowledge about infinite projections will allow us to say stuff about finite projections. The argument in Proposition 14.2.15 below feels unnecessarily convoluted, but as far as we know there is no simpler argument (compare for instance with [KR97b, Theorem 6.3.8]).

***14.2.15. Proposition.***

*Let $P, Q \in \mathcal{M}$ be two finite projections. Then $P \vee Q$ is finite.*

*Proof.* We note first that if $\mathcal{M}$ is finite, by Exercise 14.2.17 there is nothing to prove.

By Exercise 14.2.17, the projection $Q - P \wedge Q$ is finite. This means, using Kaplansky's formula (14.1) and Exercise 14.2.16, that $P \vee Q - P$ is finite. Then $P \vee Q = P + (P \vee Q - P)$ is the sum of two pairwise orthogonal finite projections. Hence we only need to prove the proposition in the case $PQ = 0$.

Let $Z$ as in Proposition 14.2.13 applied to $I_{\mathcal{M}}$. By Exercise 14.2.19 it is enough to show that $Z(P+Q)$ and $(I_{\mathcal{M}} - Z)(P+Q)$ are finite. As $Z$ is a finite projection, $Z(P+Q)$ is automatically finite. By working on $(I_{\mathcal{M}} - Z)\mathcal{M}$, we reduce to the case where $I_{\mathcal{M}}$ is properly infinite. From Remark 14.2.2 we know that any partial isometry relating $P+Q$ and its subprojections will occur in $(P+Q)\mathcal{M}(P+Q)$; so we may assume without loss of generality that $P+Q=I_{\mathcal{M}}$. Therefore the rest of the proof consists of showing that said equality is impossible.

By Proposition 14.2.14 there exists a projection $R\in\mathcal{M}$ with $I_{\mathcal{M}}\sim R\sim I_{\mathcal{M}}-R$. Applying Comparison to $P\wedge R$ and $Q\wedge(I_{\mathcal{M}}-R)$, there exists a projection $Z\in\mathcal{Z}(\mathcal{M})$ with

$$Z(P\wedge R)\preceq Z(Q\wedge(I_{\mathcal{M}}-R)),\qquad (I_{\mathcal{M}}-Z)(Q\wedge(I_{\mathcal{M}}-R))\preceq(I_{\mathcal{M}}-Z)(P\wedge R).$$

We can write (using Exercise 14.2.7 first, then Kaplansky's formula for the equivalence, and then Exercise 10.5.11)

$$\begin{aligned} ZR &= Z(R-P\wedge R)+Z(P\wedge R)\\ &\preceq Z(R-P\wedge R)+Z(Q\wedge(I_{\mathcal{M}}-R))\\ &= Z\big(R-P\wedge R+Q\wedge(I_{\mathcal{M}}-R)\big)\\ &\sim Z\big(R\vee P-P)+(I_{\mathcal{M}}-P)\wedge(I_{\mathcal{M}}-R)\big)\\ &= Z(I_{\mathcal{M}}-P)=ZQ. \end{aligned}$$

From $R\sim I_{\mathcal{M}}$ we get $Z\sim ZR\preceq ZQ$, implying that $Z$ is finite. As $\mathcal{M}$ is properly infinite, this means that $Z=0$ (being central), and therefore $Q\wedge(I_{\mathcal{M}}-R)\preceq P\wedge R$. This in turn gives us (with another use of Exercise 14.2.7 first, then Kaplansky for the equivalence, being careful that the two projections on the second line are actually orthogonal in order to apply Proposition 14.2.7, and finishing with Exercise 10.5.11)

$$\begin{aligned} I_{\mathcal{M}}-R &= \big[(I_{\mathcal{M}}-R)-Q\wedge(I_{\mathcal{M}}-R)\big]+Q\wedge(I_{\mathcal{M}}-R)\\ &\preceq \big[(I_{\mathcal{M}}-R)-Q\wedge(I_{\mathcal{M}}-R)\big]+P\wedge R\\ &\sim (I_{\mathcal{M}}-R)\vee Q-Q+P\wedge R\\ &= (I_{\mathcal{M}}-R)\vee Q+(I_{\mathcal{M}}-Q)\wedge R-Q\\ &= I_{\mathcal{M}}-Q=P. \end{aligned}$$

Then $I_{\mathcal{M}}\sim I_{\mathcal{M}}-R\preceq P$ making $I_{\mathcal{M}}$ finite, a contradiction. □

Here is a result that will be useful down the road. It should be seen as a continuity statement.

***14.2.16. Proposition.***

*Let $\{P_n\}\subset\mathcal{M}$ be an increasing sequence of finite projections, and $Q\in\mathcal{M}$ a projection with $P_n\preceq Q$ for all $n$. Then $\bigvee_n P_n\preceq Q$.*

*Proof.* We write $P = \bigvee_n P_n$. If we let $R_n = P_{n+1} - P_n$ and $R_0 = P_1$, then $P = \sum_{n=0}^{\infty} R_n$. We have $R_0 = P_1 \precsim Q$. Suppose inductively that we have pairwise orthogonal projections $S_0, \ldots, S_{n-1}$ in $\mathcal{M}$ with $S_k \sim R_k$ and $S_k \leq Q$ for all $k = 1, \ldots, n-1$. Let $Q_n = S_0 + \cdots + S_{n-1} \leq Q$. We have

$$Q_n = S_0 + \cdots + S_{n-1} \sim R_0 + \cdots + R_{n-1} = P_n.$$

By hypothesis there exists a projection $Q'_{n+1} \leq Q$ with $Q'_{n+1} \sim P_{n+1}$. If $V^*V = P_{n+1}$ and $VV^* = Q'_{n+1}$, let $Q''_n = VP_nV^*$. Then $Q''_n \leq Q'_{n+1}$ and $Q''_n \sim P_n \sim Q_n$.

As $P_n$ is finite, so is $Q_n$. Then $Q - Q_n \sim Q - Q''_n$ by Exercise 14.2.26. As $Q'_{n+1} - Q''_n \leq Q - Q''_n \sim Q - Q_n$, there exists $S_n \leq Q - Q_n$ with

$$S_n \sim Q'_{n+1} - Q''_n = VV^* - VP_nV^* \sim P_{n+1} - P_n = R_n.$$

With the full sequence $\{S_n\}$ available and using Proposition 14.2.7,

$$P = \sum_{n=0}^{\infty} R_n \sim \sum_{n=0}^{\infty} S_n \leq Q. \qquad \square$$

***14.2.17. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a $\sigma$-finite von Neumann algebra, and $P \in \mathcal{M}$ a properly infinite projection. Then $P \sim c(P)$.*

*Proof.* By Proposition 14.2.14 there exists $P_0 \leq P$ with $P \sim P_0 \sim P - P_0$. Let $V \in \mathcal{M}$ with $V^*V = P$ and $VV^* = P - P_0$; we use this partial isometry to recursively construct a sequence of projections $P_{n+1} = VP_nV^*$. As in the proof of Proposition 14.2.14, the projections $\{P_n\}$ are pairwise orthogonal and equivalent. Use Zorn's Lemma to extend the sequence $\{P_n\}$ to a maximal sequence $\{R_n\} \subset \mathcal{M}$ of pairwise orthogonal equivalent projections with $R_n \precsim P$ for all $n$; we necessarily obtain a sequence because of the $\sigma$-finiteness. By Proposition 14.2.10, $\sum_n R_n = c(P)$. Then, using Proposition 14.2.7 and the fact that $R_n \sim P_n$ for all $n$,

$$c(P) = \sum_n R_n \precsim \sum_n P_n \leq P.$$

Then $P \sim c(P)$ by Exercise 14.2.6. $\square$

***14.2.18. Corollary.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a $\sigma$-finite von Neumann algebra, and $P, Q \in \mathcal{M}$ projections with $P$ properly infinite and $c(Q) \leq c(P)$. Then $Q \precsim P$. If in addition $\mathcal{M}$ is a factor, all infinite projections are equivalent.*

*Proof.* Using Proposition 14.2.17,

$$Q \leq c(Q) \leq c(P) \sim P,$$

and so $Q \preceq P$. When $\mathcal{M}$ is a factor, $c(P) = c(Q) = I_{\mathcal{M}}$ for all nonzero projections and all infinite projections are properly infinite, and so $P \preceq Q$ and $Q \preceq P$, giving us $P \sim Q$. □

When $\mathcal{M}$ is a non-$\sigma$-finite factor, non-equivalent infinite projections may exist. To see an example, let $\mathcal{H} = \ell^2[0,1]$ (so, non-separable) and $\mathcal{M} = \mathcal{B}(\mathcal{H})$. Let $\mathcal{K}_1 \subset \mathcal{H}$ be a closed separable subspace, and $P \in \mathcal{B}(\mathcal{H})$ the orthogonal projection onto $\mathcal{K}_1$. Then $P$ is infinite, hence properly infinite (as $\mathcal{M}$ is a factor), but $P$ is not equivalent by $I_{\mathcal{H}}$ since a partial isometry $V : \mathcal{H} \to P\mathcal{H}$ would have to preserve dimension.

Next is some information about abelian projections that we will need.

**14.2.19. Proposition.**

*Let $P \in \mathcal{M}$ be an abelian projection. Then $P\mathcal{M}P = \mathcal{Z}(\mathcal{M})P$. If $Q \in \mathcal{M}$ is a projection with $Q \leq P$, then $Q = c(Q)P$.*

*Proof.* Using Corollary 14.1.4 and that $P$ is abelian,

$$P\mathcal{M}P = \mathcal{Z}(P\mathcal{M}P) = \mathcal{Z}(\mathcal{M})P.$$

As $Q \leq P$ this gives $Q = PQP$, so $Q \in \mathcal{Z}(\mathcal{M})P$. That is, $Q = TP$ with $T \in \mathcal{Z}(\mathcal{M})$. As $TP = PT$,

$$Q = \operatorname{Re} Q = \operatorname{Re}(TP) = (\operatorname{Re} T)P.$$

So we may assume that $T = T^*$. Replacing $T$ with $Tc(P)$ we may also assume that $Tc(P) = T$. We also have $0 = Q - Q^2 = (T - T^2)P$. By Corollary 14.1.4, $(T - T^2)c(P) = 0$, and so $T - T^2 = 0$. This shows that $T$ can be taken to be a central projection. Now, since $T$ is central, using Exercise 14.1.3 we get $c(Q) = Tc(P) = T$. □

**14.2.20. Proposition.**

*Let $P \in \mathcal{M}$ be a projection. The following statements are equivalent:*

*(i) $P$ is abelian;*

*(ii) $P$ is minimal (for $\preceq$) among projections with central support $c(P)$.*

*Proof.* (i) $\implies$ (ii) Suppose $Q \preceq P$ and $c(Q) = c(P)$. Then there exists $P_0 \leq P$ with $Q \sim P_0$. By Propositions 14.2.5 and 14.2.19 we get

$$Q \sim P_0 = c(P_0)P = c(P)P = P.$$

That is, $Q \sim P$ and $P$ is minimal for $\preceq$ for projections with central support $c(P)$.

(ii) $\implies$ (i) Let $Q \in P\mathcal{M}P$ be a projection. Suppose that $Q \neq c(Q)P$; since $Q \leq c(Q)P$, this means that $R = Q + (I_{\mathcal{M}} - c(Q))P$ is a proper subprojection of $P$. Noticing that $R$ has two central components and using Exercise 14.1.3,

$$c(R) = c(Q) + c\big((I_{\mathcal{M}} - c(Q))P\big) = c(Q) + c(P) - c(Q) = c(P),$$

contradicting the minimality. Hence $Q = c(Q)P$; and since von Neumann algebras are generated by their projections (Corollary 12.4.16), $P\mathcal{M}P = \mathcal{Z}(\mathcal{M})P$, abelian. □

Here is a way in which a von Neumann algebra can somehow remember its environment's dimension.

**14.2.21. Proposition.**

*Let $Z \in \mathcal{Z}(\mathcal{M})$ be a projection, and $\{P_j\}_{j\in J}$ and $\{Q_k\}_{k\in K}$ families of pairwise orthogonal abelian projections in $\mathcal{M}$ such that $c(P_j) = c(Q_k) = Z$ for all $k, j$. If*

$$\sum_{j\in J} P_j = Z = \sum_{k\in K} Q_k,$$

*then $|J| = |K|$.*

*Proof.* By working on $Z\mathcal{M}Z$ we can assume without loss of generality that $Z = I_{\mathcal{M}}$. Given $j \in J$ and $k \in K$, since both $P_j$ and $Q_k$ are abelian and $c(P_j) = c(Q_k)$ we get from Proposition 14.2.20 that $P_j \sim Q_k$. Fix $j_0 \in J$ and $k_0 \in K$. By Exercise 13.4.5

$$\mathcal{M} \simeq P_{j_0}\mathcal{M}P_{j_0} \mathbin{\overline{\otimes}} \mathcal{B}(\ell^2(J)) \simeq Q_{k_0}\mathcal{M}Q_{k_0} \mathbin{\overline{\otimes}} \mathcal{B}(\ell^2(K)).$$

We know from Exercise 14.2.15 that $P_{j_0}\mathcal{M}P_{j_0} \simeq Q_{k_0}\mathcal{M}Q_{k_0}$, and both algebras are abelian; we will denote them by $\mathcal{A}_{j_0}$ and $\mathcal{A}_{k_0}$ respectively.

Suppose first that $|J| = n < \infty$. Then $I_{\mathcal{M}} = \sum_j P_j$ is finite by Proposition 14.2.15. This forces $|K| < \infty$, for otherwise $I_{\mathcal{M}}$ would be infinite. Let $\Phi : \mathcal{A}_{j_0} \mathbin{\overline{\otimes}} \mathcal{B}(\ell^2(J)) \to \mathcal{A}_{j_0}$ be given by

$$\Phi\Big(\sum_{j_1,j_2=1}^{n} A_{j_1,j_2} \otimes E_{j_1,j_2}\Big) = \frac{1}{n}\sum_{j_1} A_{j_1,j_1}.$$

This is a "centre-valued trace" (we know from Exercise 13.4.27 that $\mathcal{Z}(\mathcal{A}_{j_0} \overline{\otimes} \mathcal{B}(\ell^2(J))) = \mathcal{A}_{j_0} \otimes I_{\mathcal{H}})$. Because $\mathcal{A}_{j_0}$ is abelian, it satisfies the property $\Phi(\tilde{A}\tilde{B}) = \Phi(\tilde{B}\tilde{A})$, with the same proof as the usual linear algebra case. We define $\Psi : \mathcal{A}_{k_0} \overline{\otimes} \mathcal{B}(\ell^2(K)) \to \mathcal{A}_{k_0}$ in the same way. The isomorphism maps $P_{j_0}$ to $P_{j_0} \otimes E_{11}$, and similarly $Q_{k_0}$ to $Q_{k_0} \otimes E_{11}$. And $\Phi(P_{j_0} \otimes E_{11}) = P_{j_0}$, a projection, and similarly $\Psi(Q_{k_0})$ is a projection. Then, considered as maps on $\mathcal{M}$,

$$n\Phi(P_{j_0}) = \sum_{j\in J} \Phi(P_j) = \Phi(I_{\mathcal{M}}) = I_{\mathcal{M}} = \Psi(I_{\mathcal{M}}) = \sum_{k\in K} \Psi(Q_k) = |K|\Psi(Q_{k_0}).$$

Hence $|K| = |J|$ since $\Phi(P_{j_0})$ and $\Psi(Q_{k_0})$ are nonzero (Exercise 14.2.21) and they are projections.

If instead one of $|J|$ and $|K|$ is infinite, both are by the previous paragraph. By Proposition 12.6.11 we have an abundance of normal states at hand. Fix one such normal state $\varphi \in \mathcal{Z}(\mathcal{M})_*$. Since $P \in \mathcal{Z}(\mathcal{M})'$, we know from Corollary 14.1.4 and Proposition 14.2.19 that there is a normal isomorphism

$$\sigma_j : \mathcal{Z}(\mathcal{M}) \to \mathcal{Z}(\mathcal{M})P_j = P_j\mathcal{M}P_j$$

given by $\sigma_j(T) = P_jTP_j$. Hence $\varphi_j(T) = \varphi \circ \sigma_j^{-1}(P_jTP_j)$ defines a normal state on $\mathcal{M}$. If $Z_0$ is the support of $\varphi$ in $\mathcal{Z}(\mathcal{M})$, the support of $\varphi_j$ is $Z_0P_j$; indeed, one checks directly that $\varphi_j(Z_0P_jT) = \varphi_j(T)$ and $\varphi_j(Z_0^\perp T) = 0$ for all $T \in \mathcal{M}$ and then the uniqueness of the support (Lemma 12.6.2) applies. For each $j \in J$, let $K_j = \{k \in K : \ \varphi_j(Q_k) \neq 0\}$. As $I_{\mathcal{M}} = \sum_k Q_k$ and $\varphi_j$ is positive and normal, $K_j$ has to be countable. When $\varphi_j(Q_k) = 0$, we have

$$0 = Z_0P_jQ_kZ_0P_j = Z_0P_jQ_k(Z_0P_jQ_k)^*,$$

so $Z_0P_jQ_k = 0$. Since $Z_0Q_k = \sum_{j\in J} Z_0P_jQ_k$, necessarily $Z_0P_jQ_k \neq 0$ for at least one $j$; that is, every $k \in K$ lies in some $K_j$. Then $K = \bigcup_{j\in J} K_j$. By Corollary 1.6.34, since each $K_j$ is countable, $|K| \leq |J|$. As the roles are symmetric, $|K| = |J|$. □

When no abelian projections are at hand, the situation is kind of the opposite and pairwise orthogonal families of projections have little cardinality restrictions:

***14.2.22. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra with no abelian projections, $n \in \mathbb{N}$, and $P \in \mathcal{M}$ a projection. Then there exist $P_1, \ldots, P_n \in \mathcal{M}$ projections, pairwise equivalent, with $P_1 + \cdots + P_n = P$.*

*Proof.* Since $P$ is not abelian, by Proposition 14.2.20 there exists a proper subprojection $P_0 \leq P$ with $c(P_0) = c(P)$. As $P - P_0 \neq 0$, $c(P - P_0) \neq 0$ and

$$c(c(P - P_0)P_0) = c(P - P_0)c(P_0) = c(P - P_0)c(P) = c(P - P_0).$$

Then we can apply Proposition 14.2.6 to $c(P-P_0)P_0$ and $P-P_0$ to get nonzero projections $P_1, P_2 \in \mathcal{M}$ with $P_1 \sim P_2$, $P_1 \leq P_0$, and $P_2 \leq P - P_0$. Since $P_1$ is not abelian, we can repeat the reasoning to get equivalent orthogonal subprojections $P_3, P_4 \leq P_1$. If $V^*V = P_1$, $VV^* = P_2$, then $VP_3V^*$ and $VP_4V^*$ are equivalent subprojections of $P_2$, also equivalent to $P_3$ and $P_4$. Hence $P_3$, $P_4$, $VP_3V^*$, and $VP_4V^*$ are pairwise equivalent and pairwise orthogonal projections. Iterating this argument we can get $2^n$ pairwise orthogonal subprojections of $P$, all equivalent. In particular we have $n$ equivalent pairwise orthogonal subprojections of $P$.

Now we apply Zorn's Lemma to the family $\mathcal{S}$ of tuples $\{\mathcal{F}_1, \ldots, \mathcal{F}_n\}$ where each $\mathcal{F}_k$ is a pairwise orthogonal family $\{Q_k^{(j)}\}_{j\in J}$ of subprojections of $P$, where $Q_1^{(j)} \sim \cdots \sim Q_n^{(j)}$ for each $j \in J$; and $\bigcup_{k=1}^n \mathcal{F}_k$ is a family of pairwise orthogonal projections. The collection $\mathcal{S}$ is nonempty by the first paragraph. We order $\mathcal{S}$ in the sense that $\{\mathcal{F}_1, \ldots, \mathcal{F}_n\} \leq \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ precisely when $\mathcal{F}_k \subset \mathcal{G}_k$ for all $k$, and the index set for the larger family contains the index set of the smaller family. Then Zorn's Lemma applies as for any chain the union of the elements (with index set the union of the index sets) will be an upper bound. So there exists $\mathcal{F} = \{\mathcal{F}_1, \ldots, \mathcal{F}_n\}$ maximal. Let, for each $k = 1, \ldots, n$, $P_k = \sum_j Q_k^{(j)}$. Then $P_k \sim P_{k'}$ for all $k, k'$ by Proposition 14.2.7. If $P - \sum_k P_k$ is nonzero, it is not abelian so we can apply the first paragraph to construct $n$ pairwise equivalent, pairwise orthogonal subprojections. With these we would contradict the maximality of $\mathcal{F}$, so $P = \sum_{k=1}^n P_k$. □

When the projection is properly infinite, we can always write it as a finite or infinite sum of projections equivalent to it (Exercise 14.2.36). If we are in a factor and we have some finite projection available to "set the size", we can do the following:

***14.2.23. Proposition.***

*Let $\mathcal{M}$ be a factor, $P, Q \in \mathcal{M}$ projections with $Q \leq P$, $P$ infinite, and $Q$ finite. Then there exist projections $\{P_j\}_{j\in J} \subset \mathcal{M}$, with $P_j \sim Q$ for all $j$ and such that $P = \sum_j P_j$. The family can be chosen so that $Q$ is one of the $P_j$.*

*Proof.* As $\mathcal{M}$ is a factor, $P$ is properly infinite. We use Zorn's Lemma to produce a maximal family of pairwise orthogonal subprojections $\{Q_j\}_{j\in J}$ of $P$, with $Q \in \{Q_j\}$ and $Q_j \sim Q$ for all $j$. Because $P$ is properly infinite, the index set $J$ is infinite Let $R_0 = P - \sum_j Q_j$. By Comparison, and since $\mathcal{M}$ is a factor, either $R_0 \preceq Q_j$ or $Q_j \preceq R_0$. This latter condition would allow us to contradict the maximality of $\{Q_j\}$ (because there is an equivalent copy of $Q$ below $R_0$), so necessarily $R_0 \preceq Q_j \sim Q$. Then $R_0$ is finite.

This means that $P - R_0$ is infinite (hence properly infinite), for otherwise $P = R_0 + (P - R_0)$ would be finite. Then we can do the following: by Exercise 14.2.31 , $R_0 \prec P - R_0$; and then

$$P - R_0 \sim (P - R_0) + R_0 = P$$

by Exercise 14.2.32. Let $V \in \mathcal{M}$ with $V^*V = P - R_0$ and $VV^* = P$. Define, for each $j \in J$, $P_j = VQ_jV^*$. Each $P_j$ is a projection because $Q_j \leq P - R_0 = V^*V$; and for the same reason they are pairwise orthogonal. We also have $P_j \sim Q_j$ via the partial isometry $VQ_j$. Indeed,

$$VQ_j(VQ_j)^* = VQ_jV^* = P_j, \qquad (VQ_j)^*VQ_j = Q_jV^*VQ_j = Q_j.$$

So $\{P_j\}$ is a family of pairwise orthogonal and mutually equivalent finite projections. Finally,

$$\sum_j P_j = \sum_j VQ_jV^* = V\Big(\sum_j Q_j\Big)V^* = V(P - R_0)V^* = VV^*VV^* = P. \qquad \square$$

The factor condition in Proposition 14.2.23 is essential. For instance if we consider $P = 1 \oplus I_{\mathcal{H}}$ as in previous examples, and $Q = 1 \oplus 0$, no other projection in $\mathbb{C} \oplus \mathcal{B}(\mathcal{H})$ is equivalent to $Q$.

We conclude with a couple results about range projections that will be useful soon. Recall that we defined the notation $[\mathcal{M}\xi]$ to denote the projection onto the subspace $\overline{\mathcal{M}\xi}$ in Definition 12.3.13, and we proved in Corollary 12.3.14 that $[\mathcal{M}\xi] \in \mathcal{M}'$.

**_14.2.24. Lemma._**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\xi, \eta \in \mathcal{H}$ cyclic vectors for $\mathcal{M}$. Then $[\mathcal{M}'\xi] \sim [\mathcal{M}'\eta]$ in $\mathcal{M}$.*

*Proof.* We know that both projections live in $\mathcal{M}$ by Corollary 12.3.14. By Lemma 12.5.11 there exist $S, T \in \mathcal{M}$ with $S \geq 0$ and $\zeta \in \overline{S\mathcal{H}}$ with $S\zeta = \xi$ and $T\zeta = \eta$. Write $P = [\mathcal{M}'\zeta]$. Since $\zeta \in P\mathcal{H}$ we have $\zeta = P\zeta$. This gives

$$P = [\mathcal{M}'\zeta] = [\mathcal{M}'P\zeta] \leq [\mathcal{M}'PS\mathcal{H}] = [PS\mathcal{M}'\mathcal{H}] \leq [PS\mathcal{H}] \leq [P\mathcal{H}] = P.$$

Using Exercise 14.2.10 and that $P, S$ are selfadjoint,

$$P = [PS\mathcal{H}] \sim [SP\mathcal{H}] = [S\mathcal{M}'\zeta] = [\mathcal{M}'S\zeta] = [\mathcal{M}'\xi].$$

This allows us to do

$$[\mathcal{M}'\eta] = [\mathcal{M}'T\zeta] = [T\mathcal{M}'\zeta] = [TP\mathcal{H}] \sim [PT\mathcal{H}] \leq P \sim [\mathcal{M}'\xi].$$

Then $[\mathcal{M}'\eta] \preceq [\mathcal{M}'\xi]$; as the roles can be exchanged, $[\mathcal{M}'\eta] \sim [\mathcal{M}'\xi]$. $\square$

***14.2.25. Lemma.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\xi, \eta \in \mathcal{H}$. Then $[\mathcal{M}'\xi] \preceq [\mathcal{M}'\eta]$ in $\mathcal{M}$ if and only if $[\mathcal{M}\xi] \preceq [\mathcal{M}\eta]$ in $\mathcal{M}'$.*

*Proof.* Suppose first that $[\mathcal{M}'\xi] \preceq [\mathcal{M}'\eta]$, so there exists $V \in \mathcal{M}$ with $V^*V = [\mathcal{M}'\xi]$ and $VV^* \leq [\mathcal{M}'\eta]$. Writing $Q = VV^*$ we have (using Exercise 14.2.11) $Q = Q[\mathcal{M}'\eta] = [Q\mathcal{M}'\eta] = [\mathcal{M}'Q\eta]$. So by working with $Q\eta$ instead of $\eta$ we may assume without loss of generality that $[\mathcal{M}'\xi] \sim [\mathcal{M}'\eta]$; for we would get $[\mathcal{M}\xi] \sim [\mathcal{M}Q\eta] \leq [\mathcal{M}\eta]$. We have

$$[\mathcal{M}'\eta] = VV^* = [V\mathcal{H}] = [VV^*V\mathcal{H}] = [V\mathcal{M}'\xi] = [\mathcal{M}'V\xi] \tag{14.3}$$

and (since $\mathcal{M}V \subset \mathcal{M}$)

$$[\mathcal{M}V\xi] \leq [\mathcal{M}\xi] = [\mathcal{M}V^*V\xi] \leq [\mathcal{M}V\xi].$$

So $[\mathcal{M}V\xi] = [\mathcal{M}\xi]$. If we show that $[\mathcal{M}\eta] \sim [\mathcal{M}V\xi]$ then we will obtain $[\mathcal{M}\eta] \sim [\mathcal{M}\xi]$ as desired.

By Proposition 14.1.3 and Eq. (14.3),

$$c([\mathcal{M}\eta]) = [\mathcal{M}'\mathcal{M}\eta] = [\mathcal{M}\mathcal{M}'\eta] = [\mathcal{M}\mathcal{M}'V\xi].$$

That is, $Z = c([\mathcal{M}\eta])$ (the central support of $[\mathcal{M}\eta]$ in $\mathcal{M}'$) is equal to $c([\mathcal{M}'\eta]) = c([\mathcal{M}'V\xi])$ (the central support of $[\mathcal{M}'\eta] = [\mathcal{M}'V\xi]$ in $\mathcal{M}$). Write $P = [\mathcal{M}'V\xi] = [\mathcal{M}'\eta] \in \mathcal{M}$. From Corollary 14.1.4 we have an isomorphism $\mathcal{M}'Z \to \mathcal{M}'P$ given by $T \longmapsto TP$. As $PV\xi$ and $P\eta$ are cyclic for $\mathcal{M}'P$, because

$$\overline{\mathcal{M}'P(PV\xi)} = \overline{\mathcal{M}'PV\xi} = P\mathcal{H}$$

and

$$\overline{\mathcal{M}'P\eta} = \overline{\mathcal{M}'\eta} = \overline{\mathcal{M}'V\xi} = P\mathcal{H},$$

from Lemma 14.2.24 (applied in $\mathcal{M}'P$) we get $[\mathcal{M}PV\xi] \sim [\mathcal{M}P\eta]$ in $\mathcal{M}'P$. But $V\xi$ and $\eta$ are in $P\mathcal{H}$, so $[\mathcal{M}V\xi] \sim [\mathcal{M}\eta]$ in $\mathcal{M}'P$, and hence in $\mathcal{M}'$.

We have thus shown that $[\mathcal{M}'\xi] \preceq [\mathcal{M}'\eta]$ in $\mathcal{M}$ implies that $[\mathcal{M}\xi] \preceq [\mathcal{M}\eta]$ in $\mathcal{M}'$. The reverse implication is now trivial by using that $\mathcal{M} = (\mathcal{M}')'$. $\square$

**14.2.1. Exercises.**

(14.2.1) Let $P \in \mathcal{M}$ be a projection with $P \preceq 0$. Show that $P = 0$.

(14.2.2) Let $P, Q \in M_n(\mathbb{C})$ be projections. Show that $P \sim Q$ if and only if $\operatorname{Tr}(P) = \operatorname{Tr}(Q)$.

(14.2.3) Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra and $P, Q \in \mathcal{M}$ be projections. Show that $P \sim Q$ if and only if $\operatorname{Tr}(ZP) = \operatorname{Tr}(ZQ)$ for every central projection $Z$.

(14.2.4) Let $P, Q, R \in \mathcal{M}$ be projections, with $P \preceq Q$ and $Q \preceq R$. Show that $P \preceq R$.

(14.2.5) Let $P, Q, Z \in \mathcal{M}$ be projections, with $Z \in \mathcal{Z}(\mathcal{M})$ and such that $P \preceq Q$. Show that $ZP \preceq ZQ$.

(14.2.6) Let $P, Q \in \mathcal{M}$ with $P \preceq Q$ and $Q \preceq P$. Prove that $P \sim Q$ by structuring the argument the same as the proof of Schröder–Bernstein (Theorem 1.6.13).

(14.2.7) Let $\mathcal{M}$ be a von Neumann algebra and $\{P_j\}$ and $\{Q_j\}$ two families of pairwise orthogonal projections in $\mathcal{M}$, such that $P_j \preceq Q_j$ for all $j$. Show that $\sum_j P_j \preceq \sum_j Q_j$.

(14.2.8) Show that if in Proposition 14.2.16 $\mathcal{M}$ is $\sigma$-finite and $Q$ is properly infinite, the result holds without requiring the $P_n$ to be finite.

(14.2.9) Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space. Let $P, Q \in \mathcal{B}(\mathcal{H})$ be infinite projections. Show that $P \sim Q$. Is $I_{\mathcal{H}} - Q \sim I_{\mathcal{H}} - Q$?

(14.2.10) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $T \in \mathcal{M}$. Show that $[T\mathcal{H}] \sim [|T|\mathcal{H}] = [T^*\mathcal{H}]$ in $\mathcal{M}$.

(14.2.11) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $P \in \mathcal{M}$ a projection, and $\xi \in \mathcal{H}$. Show that $[P\mathcal{M}\xi] = P[\mathcal{M}\xi]$.

(14.2.12) Let $\mathcal{M}$ be a von Neumann algebra, and $P \in \mathcal{M}$ a projection. Show that $P$ is minimal if and only if $Q \leq P$ for a projection $Q \in \mathcal{M}$, implies that either $Q = P$ or $Q = 0$.

(14.2.13) Let $\mathcal{M} = M_n(\mathbb{C})$. Find all the minimal projections. Find all the abelian projections. Show that $\mathcal{M}$ is finite.

(14.2.14) Let $\{P_j\} \subset \mathcal{M}$ be a family of pairwise equivalent projections. Fix a projection $P \in \mathcal{M}$ with $P \sim P_j$ for all $j$. Show that

$$c\Big(\bigvee_j P_j\Big) = c(P).$$

(14.2.15) Let $P, Q \in \mathcal{M}$ be equivalent. Show that $P\mathcal{M}P \simeq Q\mathcal{M}Q$.

(14.2.16) Let $P, Q \in \mathcal{M}$ be equivalent projections. Show that if $P$ is any of minimal, abelian, infinite, finite, purely infinite, properly infinite, then so is $Q$.

(14.2.17) Let $P, Q \in \mathcal{M}$ be projections with $Q \preceq P$. Show that if $P$ is finite, then $Q$ is finite.

(14.2.18) Let $P, Q \in \mathcal{M}$ be projections with $Q \preceq P$. Show that if $P$ is abelian, then $Q$ is abelian.

(14.2.19) Let $P, Z \in \mathcal{M}$ be projections with $Z$ central. Show, without using Proposition 14.2.15, that $P$ is finite if and only if $ZP$ and $(I_{\mathcal{M}} - Z)P$ are finite.

(14.2.20) Let $\mathcal{M}$ be a finite von Neumann algebra. Show that the only isometries are the unitaries.

(14.2.21) Let $\Phi$ as in the proof of Proposition 14.2.21. Show that it is faithful.

(14.2.22) Let $\mathcal{H}$ be an infinite-dimensional separable Hilbert space. Fix an orthonormal basis and consider the associated unilateral shift $S$. Show that $S$ is a partial isometry, and conclude that $\mathcal{B}(\mathcal{H})$ is infinite.

(14.2.23) Show an example of a infinite projection that cannot be halved.

(14.2.24) Let $\mathcal{M}$ be a finite von Neumann algebra, and $P, Q \in \mathcal{M}$ projections with $P \sim Q$. Show that $I_{\mathcal{M}} - P \sim I_{\mathcal{M}} - Q$.

(14.2.25) Let $\mathcal{M}$ be a finite von Neumann algebra, and $P, Q \in \mathcal{M}$ projections with $P \preceq Q$. Show that $I_{\mathcal{M}} - Q \preceq I_{\mathcal{M}} - P$.

(14.2.26) Let $\mathcal{M}$ be a von Neumann algebra, and $P, Q \in \mathcal{M}$ finite projections with $P \sim Q$. Show that $I_{\mathcal{M}} - P \sim I_{\mathcal{M}} - Q$.

(14.2.27) Let $P \in \mathcal{M}$ be a finite projection, and $Q \in \mathcal{M}$ a projection. Show that there exists a number $s \in \mathbb{N} \cup \{0\}$ such that $s$ is the maximum such that there exist pairwise orthogonal projections $\{P_1, \ldots, P_s\} \subset \mathcal{M}$ with $P_k \sim Q$ for all $k$, and $\sum_{k=1}^{s} P_k \preceq P$.

(14.2.28) Let $\mathcal{M}$ be a finite von Neumann algebra and $T \in \mathcal{M}$ with polar decomposition $T = V|T|$. Show that $V$ can be extended to a unitary $U$ with $T = U|T|$.

(14.2.29) Let $\mathcal{M}$ be a finite-dimensional von Neumann algebra. Show that $\mathcal{M}$ is finite, in two ways:

(i) by using the explicit form of a finite-dimensional von Neumann algebra;

(ii) by a direct argument.

(14.2.30) Let $\mathcal{M}$ be a von Neumann algebra. Show that the following statements are equivalent:

(i) $\mathcal{M}$ is finite;

(ii) for any projections $P, Q \in \mathcal{M}$ with $P \sim Q$, there exists $U \in \mathcal{M}$ unitary with $Q = UPU^*$.

(14.2.31) Let $\mathcal{M}$ be a factor, and $P, Q \in \mathcal{M}$ projections with $Q$ finite and $P$ infinite. Show that $Q \prec P$.

(14.2.32) Let $P, Q \in \mathcal{M}$ be projections with $P$ properly infinite and $Q \preceq P$. Show that $P \vee Q \sim P$.

(14.2.33) Let $P, Q \in \mathcal{M}$ be properly infinite, with $P + Q = I_{\mathcal{M}}$ and $P \sim Q$. Show that $P \sim I_{\mathcal{M}}$.

(14.2.34) Let $P_1, P_2, Q_1, Q_2 \in \mathcal{M}$ be projections with $P_1 + P_2 = Q_1 + Q_2$, $P_1 P_2 = 0 = Q_1 Q_2$, and $P_1 \sim P_2$, $Q_1 \sim Q_2$. Show that $P_1 \sim Q_1$.

(14.2.35) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra, $Q \in \mathcal{M}$ a projection, and $\mathcal{K}$ a Hilbert space. Show that $c(Q \otimes E_{11}) = c(Q) \otimes I_{\mathcal{K}}$ in $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$.

(14.2.36) Let $P \in \mathcal{M}$ be a properly infinite projection, and $n \in \mathbb{N} \cup \{\infty\}$. Show that there exist pairwise orthogonal projections $\{P_k\}_{k=1}^n \subset \mathcal{M}$ with $P_k \sim P$ for all $k$, and $P = \sum_k P_k$.

## 14.3. Classification of von Neumann Algebras

von Neumann algebras are classified in types according to the behaviour of their projections. Mixed types exist, as we will discuss.

**14.3.1. Definition.**

*A von Neumann algebra $\mathcal{M}$ is of **type***

- *I, if every projection in $\mathcal{M}$ has a nonzero abelian subprojection;*
    - *$\mathrm{I}_n$, for $n \in \mathbb{N} \cup \{\infty\}$, if $\mathcal{M}$ admits a family $\{E_j\}_{j=1}^n$ of pairwise equivalent abelian projections with $c(E_j) = I_{\mathcal{M}}$ for all $j$;*
- *II, if $\mathcal{M}$ has no abelian projections, and every projection in $\mathcal{M}$ has a nonzero finite subprojection;*
    - *$\mathrm{II}_1$, if $\mathcal{M}$ is of type II and $I_{\mathcal{M}}$ is finite;*
    - *$\mathrm{II}_\infty$, if $\mathcal{M}$ is of type II and $I_{\mathcal{M}}$ is infinite;*
- *III, if $\mathcal{M}$ has no finite projections.*

*The subscripted versions are pronounced "one–enn" (or "one–five", etc.), "one–infinity", "two–one", "two–infinity".*

The definition of type $\mathrm{I}_n$ makes sense due to Proposition 14.2.21.

For some basic examples, $M_n(\mathbb{C})$ is of type $\mathrm{I}_n$, $\mathcal{B}(\mathcal{H})$ is of type $\mathrm{I}_\infty$ when $\dim \mathcal{H} = \infty$. And $L(G)$, as in Example 13.3.13, is of type $\mathrm{II}_1$; this will be trivial to see after the Classification Theorem. A von Neumann algebra $\mathcal{M}$ does not have to be of one of the types. For instance we can form $M_2(\mathbb{C}) \oplus L(G)$, acting on $\mathbb{C}^2 \oplus \ell^2(G)$; it has one block of type $\mathrm{I}_2$ and another of type $\mathrm{II}_1$. More algebraically, the central projections $P_1 = I_2 \oplus 0$ and $P_2 = 0 \oplus I_G$ satisfy that $P_1\mathcal{M}$ is of type $\mathrm{I}_2$ and $P_2\mathcal{M}$ is of type $\mathrm{II}_1$. This block-like decomposition is always possible:

**14.3.2. Theorem. (Type Classification)**

*Let $\mathcal{M}$ be a von Neumann algebra. Then there exist unique central projections $P_I$, $P_{II}$, and $P_{III}$, that add to $I_{\mathcal{M}}$, and such that $P_I\mathcal{M}$ is of type I, $P_{II}\mathcal{M}$ is of type II, and $P_{III}\mathcal{M}$ is of type III.*

*The projection $P_I$ in turn subdivides as $P_I = \sum_{n\leq \dim\mathcal{H}} P_n$, with each $P_n$ central and $P_n\mathcal{M}$ of type $I_n$. The projection $P_{II}$ can be written $P_{II} = P_{II_1} + P_{II_\infty}$, both central and $P_{II_1}\mathcal{M}$ of type $II_1$, and $P_{II_\infty}\mathcal{M}$ of type $II_\infty$.*

*Proof.* Let $P_I = \bigvee\{Q \in \mathcal{M},\ \text{abelian}\}$. Then $P_I \in \mathcal{M}$ by Corollary 12.3.9. If $Q$ is abelian, then $UQU^*$ is abelian for any unitary (Exercise 14.2.16). By Exercise 10.5.9, $UP_IU^* = P_I$ for all unitaries $U \in \mathcal{M}$. This means that $P_I \in \mathcal{M}'$, and hence $P_I \in \mathcal{Z}(\mathcal{M})$. We claim that $P_I\mathcal{M}$ is of type I. Let $P \in P_I\mathcal{M}$ be a nonzero projection. By construction of $P_I$, since $P = PP_I$ there exists $Q \in \mathcal{M}$, abelian, with $QP \neq 0$. By Proposition 14.2.6, this implies that there exist nonzero subprojections $P_0 \leq P$ and $Q_0 \leq Q$ with $P_0 \sim Q_0$. As $Q_0$ is abelian (because $Q_0\mathcal{M}Q_0 \subset Q\mathcal{M}Q$), $P_0$ is abelian by Exercise 14.2.16 and hence $P_I\,\mathcal{M}$ is type I.

Now we work on $(I_{\mathcal{M}} - P_I)\mathcal{M}$, which has no nonzero abelian projections. Let $P_{II} = \bigvee\{Q \in \mathcal{M},\ \text{finite}\}$. It is in $\mathcal{M}$ by Corollary 12.3.9. Given a unitary $U \in \mathcal{M}$, $Q$ is finite if and only if $UQU^*$ is finite. Reasoning as in the previous paragraph, $P_{II} = UP_{II}U^*$, and we conclude that $P_{II} \in \mathcal{Z}(\mathcal{M})$. Given a projection $P \in P_{II}\mathcal{M}$, as in the previous paragraph we find $Q \in \mathcal{M}$, finite, and nonzero subprojections $P_0 \leq P$ and $Q_0 \leq Q$ with $P_0 \sim Q_0$. As $Q_0$ is finite (Exercise 14.2.17), by Exercise 14.2.16 we have that $P_0$ is finite. This shows that $P_{II}\mathcal{M}$ is of type II.

Finally we define $P_{III} = I_{\mathcal{M}} - P_I - P_{II}$. By construction, $P_{III}\mathcal{M}$ has no finite projections and is hence of type III.

Regarding the uniqueness, suppose that $Q_I$, $Q_{II}$, $Q_{III}$ are pairwise orthogonal central projections that compress $\mathcal{M}$ to the respective types. The algebra $P_IQ_{II}\mathcal{M}$ has no abelian projections (because $Q_{II}\mathcal{M}$ hasn't any). But if $P_IQ_{II}$ is nonzero, being a subprojection of $P_I$ it has to have a nonzero abelian subprojection. So $P_IQ_{II} = 0$. Similarly, $P_IQ_{III} = 0$. Thus $P_I = Q_I$. We can reason similarly with $P_{II}Q_{III}$: if nonzero, it both has and cannot have finite subprojections. So $P_{II} = Q_{II}$ and $P_{III} = Q_{III}$.

Fix $n \leq \dim\mathcal{H}$ a cardinal (either a natural number or an infinity less than or equal $\dim\mathcal{H}$; when $\mathcal{H}$ is separable, $n \in \mathbb{N}\cup\{\infty\}$). Let $\{Q_j\}$ be a maximal family of pairwise orthogonal central projections in $\mathcal{M}$, with each $Q_j = \sum_{k=1}^n Q_{j,k}$ where $Q_{j,k} \sim Q_{j,h}$ and $Q_{j,k}$ abelian for all $k, h, j$. From Proposition 14.2.10 we know that $Q_j = c(Q_j) = c(Q_{j,k})$ for all $k$. Then $R_k = \sum_j Q_{j,k}$ is an abelian projection and $c(R_k) = \sum_j c(Q_{j,k}) = \sum_j Q_j$.

If we put $P_n = \sum_j Q_j$, then $P_n$ is a central projection and $P_n = \sum_{k=1}^n R_k$, equivalent abelian projections. So either $P_n = 0$ or $P_n\mathcal{M}$ is of type $\mathrm{I}_n$. As $P_n\mathcal{M}$ has abelian projections if nonzero, $P_n \leq P_I$. Having done this for every $n$, let $P = P_I - \sum_n P_n$. Suppose that $P \neq 0$. By Exercise 14.3.2 there exists a projection $R_0 \in P_I\mathcal{M}$, abelian, with $c(R_0) = P_I$. As $P$ is central, $c(R_0P) = P$. Let $\{R_j\}_{j\in J}$ be a maximal family of pairwise orthogonal projections in $P\mathcal{M}$, with $R_j \sim R_0$ for all $j$. By Comparison there exists a nonzero projection $Z \in \mathcal{Z}(P\mathcal{M})$ with $Z(P - \sum_j R_j) \prec ZR_0P$ ($Z$ is nonzero for otherwise we get $R_0P \preceq P - \sum_j R_j$, contradicting the maximality). So there exists a projection $R_0' \in P\mathcal{M}$ such that $Z - \sum_j ZR_j \sim R_0' \leq ZR_0$, with $R_0' \neq ZR_0$. We have $c(R_0') = c\big(Z - \sum_j ZR_j\big)$. As $R_0' \leq ZR_0$ and $ZR_0$ is abelian, $R_0'$ is abelian. If $R_0' = 0$ then $Z$ is a sum of equivalent abelian projections; as $Z \leq P$, this contradicts how $P_n$, for $n = |J|$, was constructed. If instead $R_0' \neq 0$, then for each $j$ there exists $R_j' \leq R_j$ with $R_j' \sim R_0'$ (so, abelian) and $c(R_j') = c(R_0') \leq P$. Then $\{R_j'\}$ are pairwise equivalent abelian projections below $P$; again a contradiction. It follows that $P = 0$, which is to say that $P_I = \sum_{n \leq \dim \mathcal{H}} P_n$.

Applying Proposition 14.2.13 to $P_{II}$ we get central projections $P_{II_1}$ and $P_{II_\infty}$ with $P_{II_1}$ finite, $P_{II_\infty}$ properly infinite, and $P_{II} = P_{II_1} + P_{II}$. The algebra $P_{II_1}\mathcal{M}$ is of type $\mathrm{II}_1$ since it is type II (it cannot have abelian projections, and every projection is above a nonzero finite projection) and its unit $P_{II_1}$ is finite. And $P_{II_\infty}\mathcal{M}$ is of type $\mathrm{II}_\infty$ since it is of type II and its unit is properly infinite. $\square$

### *14.3.3. Corollary.*

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a factor. Then $\mathcal{M}$ is of a fixed type $I_n$ (with $n \leq \dim \mathcal{H}$), $II_1$, $II_\infty$, or III.*

*Proof.* The centre of $\mathcal{M}$ is trivial, so only one of the projections in Theorem 14.3.2 is nonzero. $\square$

Type decomposition plays nicely with compressions and amplifications, in the following sense.

### *14.3.4. Proposition.*

*Let $\mathcal{M}$ be a von Neumann algebra of fixed type I, II, or III. Let $P \in \mathcal{M}$ be a projection, and $\mathcal{H}$ a Hilbert space. Then $P\mathcal{M}P$ and $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$ have the same type as $\mathcal{M}$.*

*Proof.* We will do the compressions first. Suppose that $\mathcal{M}$ is of type I. Let $Q \in P\mathcal{M}P$ be a nonzero projection. Then $Q \leq P$ in $\mathcal{M}$. By the assumption and Exercise 14.3.2 there exists a projection $R \in \mathcal{M}$, abelian, with $c(R) = I_\mathcal{M}$. By Proposition 14.2.6 there exist nonzero projections $Q_0 \leq Q$ and $R_0 \leq R$ with $Q_0 \sim R_0$. As $R$ is abelian, so is $R_0$; then $Q_0$ is abelian by Exercise 14.2.16. Hence $P\mathcal{M}P$ is type I.

If $\mathcal{M}$ is type II, let $Q \in P\mathcal{M}P$ be a nonzero projection. As $Q$ is also a projection in $\mathcal{M}$, it cannot have an abelian subprojection, and it does have a finite proper subprojection $Q_0$. And $Q_0 = QQ_0Q = PQQ_0QP \in P\mathcal{M}P$. Hence every projection in $P\mathcal{M}P$ has a finite proper subprojection, and thus $P\mathcal{M}P$ is type II.

If $\mathcal{M}$ is type III, it has no finite projections. Suppose that $Q \in P\mathcal{M}P$ is finite in $P\mathcal{M}P$. If $Q_0 \leq Q$ in $\mathcal{M}$ with $Q_0 = V^*V$ and $Q = VV^*$ for $V \in \mathcal{M}$, by Exercise 10.4.12 we have $V = QVQ_0 = PQVQ_0P \in P\mathcal{M}P$, so $Q_0 \sim Q$ in $P\mathcal{M}P$; this forces $Q_0 = Q$, and $Q$ is finite in $\mathcal{M}$, a contradiction. Hence $P\mathcal{M}P$ has no finite projections and is thus of type III.

We know from Exercise 13.4.27 that $\mathcal{Z}(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H})) = \mathcal{Z}(\mathcal{M}) \otimes I_\mathcal{H}$. When $\mathcal{M}$ is type II, consider the central projection $\tilde{P}_I$ such that $\tilde{P}_I(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H}))$ is type I; we have $\tilde{P}_I = P \otimes I_\mathcal{H}$ with $P \in \mathcal{Z}(\mathcal{M})$, and then $P\mathcal{M} \simeq (P \otimes E_{11})(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H}))(P \otimes E_{11})$ (Exercise 14.1.5) is type I as it is a corner of a type I. But by hypothesis $\mathcal{M}$ has no type I part, so $P = 0$ and hence $\tilde{P}_I = 0$. This reasoning applies to any combination of distinct types, so if $\mathcal{M}$ is of a fixed type, then $\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{H})$ is of the same fixed type. $\square$

The results of Proposition 14.3.4 do not extend to the more granular cases $\mathrm{I}_n$ and $\mathrm{II}_1$, $\mathrm{II}_\infty$. For instance $M_5(\mathbb{C})$ is type $\mathrm{I}_5$ but the corner determined by a rank-2 projection is type $\mathrm{I}_2$. And $M_5(\mathbb{C})\,\overline{\otimes}\,\mathcal{B}(\ell^2(\mathbb{N}))$ is of type $\mathrm{I}_\infty$. Similarly, if $\mathcal{M}$ is a $\mathrm{II}_1$-factor then $\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\ell^2(\mathbb{N}))$ is a type $\mathrm{II}_\infty$ factor, and this also shows that a corner of a $\mathrm{II}_\infty$-factor can be a $\mathrm{II}_1$-factor.

**14.3.1. Type I von Neumann Algebras.** With the classification at hand, we can get a full understanding of what a type I von Neumann algebra looks like.

***14.3.5. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a type I von Neumann algebra. Then there exist pairwise orthogonal projections $\{P_k\}_{k\in K}$, $K \subset \{1,\ldots,\dim\mathcal{H}\}$, that add to $I_\mathcal{M}$ and such that $P_k\mathcal{M} \simeq \mathcal{B}(\mathcal{H}_k)\,\overline{\otimes}\,\mathcal{A}_n$, where $\dim\mathcal{H}_k = k$ and $\mathcal{A}_k$ is an abelian von Neumann algebra. Equivalently,*

$$\mathcal{M} \simeq \bigoplus_{k\in K} \mathcal{B}(\mathcal{H}_k)\,\overline{\otimes}\,\mathcal{A}_k. \tag{14.4}$$

*Proof.* By Theorem 14.3.2 we just have to show that if $\mathcal{M}$ is type $\mathrm{I}_n$, then $\mathcal{M} \simeq \mathcal{B}(\mathcal{H}_1) \mathbin{\overline{\otimes}} \mathcal{A}$, with $\dim \mathcal{H}_1 = n$, for an abelian von Neumann algebra $\mathcal{A}$. By hypothesis there exist pairwise equivalent abelian projections $\{Q_1, \ldots, Q_n\} \subset \mathcal{M}$ ($n$ is possibly infinite), that add to $I_{\mathcal{M}}$, and with $c(Q_h) = I_{\mathcal{M}}$ for all $h$. So we have partial isometries $E_{11}, E_{12}, \ldots$ such that

$$E_{1h}^* E_{1h} = Q_h, \qquad E_{1h} E_{1h}^* = Q_1.$$

Then we define

$$E_{hh} = Q_h, \qquad E_{kj} = E_{1k}^* E_{1j}, \tag{14.5}$$

which satisfy the matrix unit relations $E_{kj} E_{ab} = \delta_{j,a} E_{kb}$ and $E_{kj}^* = E_{jk}$ (Exercise 14.3.1). Let $\mathcal{A} = Q_1 \mathcal{M} Q_1 \subset \mathcal{B}(Q_1 \mathcal{H})$, which is abelian by hypothesis. Let $\mathcal{H}_1 = \ell^2(\{1, \ldots, n\})$ and $\mathcal{H}_2 = Q_1 \mathcal{H}$. Let $U : \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}$ be the linear map induced by, for $J$ finite,

$$U\Big(\sum_{j\in J} \delta_j \otimes \xi_j\Big) = \sum_{j\in J} E_{j1} \xi_j.$$

This gives us

$$\begin{aligned}
\Big\langle U\Big(\sum_{j\in J} \delta_j \otimes \xi_j\Big), U\Big(\sum_{j\in J} \delta_j \otimes \eta_j\Big)\Big\rangle &= \Big\langle \sum_{j\in J} E_{j1}\xi_j, \sum_{j\in J} E_{j1}\eta_j \Big\rangle \\
&= \sum_{k,j\in J} \langle E_{k1}^* E_{j1} \xi_j, \eta_j\rangle = \sum_{j\in J} \langle \xi_j, \eta_j\rangle \\
&= \Big\langle \sum_{j\in J} \delta_j \otimes \xi_j, \sum_{j\in J} \delta_j \otimes \eta_j \Big\rangle.
\end{aligned}$$

Hence $U$ is isometric on the algebraic tensor product, so it extends to an isometry $\mathcal{H}_1 \mathbin{\overline{\otimes}} \mathcal{H}_2 \to \mathcal{H}$. We also have that $U$ is surjective, since its range is closed and it contains $Q_j \mathcal{H}$ for all $j$. That is, $U$ is a unitary. Let $\{F_{rs}\}$ be the canonical matrix units for $\mathcal{H}_1$. Elements of the form $X = \sum_{r,s\in R} F_{rs} \otimes A_{rs}$, where the set $R$ of indices is finite, are sot-dense in $\mathcal{B}(\mathcal{H}_1) \mathbin{\overline{\otimes}} \mathcal{A}$. We have, for $\xi \in \mathcal{H}$,

$$\begin{aligned}
UXU^*\xi = UXU^*\Big(\sum_{j=1}^n E_{j1}E_{1j}\xi\Big) &= UX\Big(\sum_{j=1}^n \delta_j \otimes E_{1j}\xi\Big) \\
&= U\Big(\sum_{r,s\in R} \sum_{j=1}^n F_{rs}\delta_j \otimes A_{rs}E_{1j}\xi\Big) \\
&= U\Big(\sum_{r,s\in R} \delta_r \otimes A_{rs}E_{1s}\xi\Big) = \sum_{r\in R}\sum_{s\in R} E_{r1}A_{rs}E_{1s}\xi.
\end{aligned}$$

As $A_{rs} \in E_{11}\mathcal{M}E_{11} = Q_1\mathcal{M}Q_1$ we get that

$$UXU^* = \sum_{r\in R}\sum_{s\in R} E_{r1}A_{rs}E_{1s} \in \mathcal{M}.$$

If now $T \in \mathcal{M}$,

$$\begin{aligned} U^*TU\Big(\sum_j \delta_j \otimes \xi_j\Big) &= U^* \sum_j TE_{j1}\xi_j = U^* \sum_k E_{k1} \sum_j E_{1k}TE_{j1}\xi_j \\ &= \sum_{k,j} \delta_k \otimes E_{1k}TE_{j1}\xi_j \\ &= \Big(\sum_{r,s} F_{rs} \otimes E_{1r}TE_{s1}\Big)\Big(\sum_j \delta_j \otimes \xi_j\Big) \end{aligned}$$

Therefore $U^*TU \in \mathcal{B}(\mathcal{H}_1)\,\overline{\otimes}\,\mathcal{A}$, since $E_{1r}TE_{s1} \in \mathcal{A}$ for all $r, s$. We have shown that the map $T \longmapsto U^*TU$ is a $*$-monomorphism with wot-dense range, and $\sigma$-weakly continuous by Exercise 12.6.4. By Corollary 12.6.14, it is surjective. Hence $\gamma : \mathcal{B}(\mathcal{H}_1)\,\overline{\otimes}\,\mathcal{A} \to \mathcal{M}$ given by $\gamma(X) = UXU^*$ is a $*$-isomorphism. $\square$

When $\mathcal{M}$ is a type I factor, there can only be one summand in (14.4). So $\mathcal{M} = \mathcal{B}(\mathcal{H})\,\overline{\otimes}\,\mathcal{A}$ for some Hilbert space $\mathcal{H}$ and $\mathcal{A}$ abelian. As elements $I_{\mathcal{H}} \otimes A$ for $A \in \mathcal{A}$ are central, $\mathcal{M}$ being a factor forces $\mathcal{A}$ has to be trivial, $\mathcal{A} = \mathbb{C}\, I_{\mathcal{A}}$. So $\mathcal{M} \simeq \mathcal{B}(\mathcal{H})$. If $\mathcal{M}$ is type $\mathrm{I}_n$ with $n \in \mathbb{N}$, then $\dim \mathcal{H} = n$ and $\mathcal{M} \simeq M_n(\mathbb{C})$. When $n \in \mathbb{N}$ in the proof of Proposition 14.3.5, $\mathcal{B}(\mathcal{H}_k)\,\overline{\otimes}\,\mathcal{A}_k \simeq M_k(\mathcal{A}_k)$.

We will consider below a particular case of a type I algebra. We defined the notion of two projections being in generic position in Definition 10.5.8.

***14.3.6. Proposition.***

*Let $P, Q \in \mathcal{B}(\mathcal{H})$ be two projections. Then $\mathcal{M} = W^*(P, Q)$ is a von Neumann algebra of type I. More concretely, there exists a central projection $Z \in \mathcal{Z}(\mathcal{M})$ with $Z\mathcal{M}$ of type $I_2$ and $(I_{\mathcal{M}} - Z)\mathcal{M}$ abelian of dimension at most* 4.

*Proof.* For space reasons in the formulas, we will temporarily used the notation $P^\perp = I_{\mathcal{M}} - P$. Let

$$P_0 = P - P \wedge Q - P \wedge Q^\perp, \qquad Q_0 = Q - P \wedge Q - P^\perp \wedge Q.$$

We claim that

$$I_{\mathcal{M}} = P \wedge Q + P \wedge Q^\perp + P^\perp \wedge Q + P^\perp \wedge Q^\perp + P_0 \vee Q_0. \tag{14.6}$$

Indeed, we have

$$\begin{aligned} I_{\mathcal{M}} - P_0 \vee Q_0 &= P_0^\perp \wedge Q_0^\perp \\ &= (P^\perp + P \wedge Q + P \wedge Q^\perp) \wedge (Q^\perp + P \wedge Q + P^\perp \wedge Q). \end{aligned}$$

Each group of three projections inside the brackets is pairwise orthogonal. So if $\xi \in (P_0^\perp \wedge Q_0^\perp)\mathcal{H}$, we have

$$P^\perp\xi = P^\perp(Q^\perp + P \wedge Q + P^\perp \wedge Q)\xi = P^\perp Q^\perp \xi + (P^\perp \wedge Q)\xi$$

$$= P^\perp Q^\perp (P^\perp + P\wedge Q + P\wedge Q^\perp)\xi + (P^\perp\wedge Q)\xi$$
$$= P^\perp Q^\perp P^\perp \xi + (P^\perp \wedge Q)\xi.$$

Iterating this, $P^\perp\xi = (P^\perp Q^\perp P^\perp)^n\xi + (P^\perp\wedge Q)\xi$ for all $n\in\mathbb{N}$. By Proposition 12.1.17, $P^\perp\xi = (P^\perp\wedge Q^\perp)\xi + (P^\perp\wedge Q)\xi$. With a similar computation,

$$P\xi = (P\wedge Q)\xi + (P\wedge Q^\perp)\xi \tag{14.7}$$

(this equality does not hold for arbitrary projections, see Exercise 10.5.8). It follows that

$$I_{\mathcal{M}} - P_0\vee Q_0 = P\wedge Q + P\wedge Q^\perp + P^\perp\wedge Q + P^\perp\wedge Q^\perp,$$

which is (14.6). Now we have that $P_0$ and $Q_0$ are in generic position in $(P_0\vee Q_0)\mathcal{H}$:

$$P_0\wedge Q_0 = P_0\wedge Q_0^\perp = P_0^\perp\wedge Q_0 = P_0^\perp\wedge Q_0^\perp = 0$$

(Exercise 14.3.3). Since $\mathcal{M}$ is generated by $P$ and $Q$ and $P(P\wedge Q) = (P\wedge Q)P$ and $Q(P\wedge Q) = (P\wedge Q)Q$, we get that $P\wedge Q\in\mathcal{Z}(\mathcal{M})$, since it commutes with the generators. Similarly, $P\wedge Q^\perp, P^\perp\wedge Q, P^\perp\wedge Q^\perp\in\mathcal{Z}(\mathcal{M})$. Let us denote these four central projections $R_1, R_2, R_3, R_4$. It follows from (14.6) that $P_0\wedge Q_0 = I_{\mathcal{M}} - R_1 - R_2 - R_3 - R_4 \in \mathcal{Z}(\mathcal{M})$. As each $R_jP, R_jQ\in\{0, R_j\}$, $R_j\mathcal{M} = \mathbb{C}R_j$. With $Z = P_0\vee Q_0\in\mathcal{Z}(\mathcal{M})$, we have that

$$\mathcal{M} = Z\mathcal{M} + \sum_{j=1}^4 \mathbb{C}R_j.$$

So $(I_{\mathcal{M}} - Z)\mathcal{M} = \sum_{j=1}^4\mathbb{C}R_j$ is abelian with dimension between 0 and 4, depending on whether each $R_j$ is nonzero or not. By Proposition 10.5.12 we can think of $P_0$ and $Q_0$ in $Z\mathcal{M}$ as

$$P_0 = \begin{bmatrix} I & 0\\ 0 & 0\end{bmatrix}, \qquad Q_0 = \begin{bmatrix} C^2 & CS\\ CS & S^2\end{bmatrix},$$

where $C, S$ are positive commuting contractions with $C^2+S^2 = I$. The proof will be finished if we show that $Z\mathcal{M} = W^*(P_0, Q_0) = M_2(P_0\mathcal{M}P_0)$, and that $P_0\mathcal{M}P_0$ is abelian.

As $Z$ is central and $\mathcal{M} = W^*(P,Q)$, we have $Z\mathcal{M} = W^*(ZP, ZQ) = W^*(P_0, Q_0)$. Also,

$$P_0Q_0P_0 = \begin{bmatrix} C^2 & 0\\ 0 & 0\end{bmatrix}, \tag{14.8}$$

so $C = (C^2)^{1/2}\in P_0\mathcal{M}P_0$; this shows that $W^*(C)\subset P_0\mathcal{M}P_0$. The algebra $P_0\mathcal{M}P_0$ is generated by words of the form $P_0wP_0$, where $w$ is a word on $P, Q$. As $P_0 = PZ$ and $Q_0 = QZ$, we may assume that $w$ is a word on $P_0, Q_0$, and (14.8) shows that $P_0wP_0 = P_0C^mP_0$ for some $m$. It follows that $P_0\mathcal{M}P_0 = W^*(C)$, abelian.

Suppose that $CS\xi = 0$. Then

$$Q_0 \begin{bmatrix} \xi \\ 0 \end{bmatrix} = \begin{bmatrix} C^2 & CS \\ CS & S^2 \end{bmatrix} \begin{bmatrix} \xi \\ 0 \end{bmatrix} = \begin{bmatrix} C^2\xi \\ 0 \end{bmatrix} \in P_0\mathcal{H},$$

and if $\xi \neq 0$ this would make $P_0 \wedge Q_0 \neq 0$. So $CS = C(1-C^2)^{1/2}$ is injective, and hence $C$ and $S$ are injective; this means that both, being also positive, have dense range. Within $Z\mathcal{M}$ we can do manipulations like this: for any $X \in W^*(C)$,

$$\begin{bmatrix} X & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C^2 & CS \\ CS & S^2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & XCS \\ 0 & 0 \end{bmatrix}.$$

Using $X = 1_{[0,1-\delta]}(C)\left((\varepsilon I_n + S\right)^{-1}$ and then taking sot limit as $\varepsilon \to 0$ (and remembering that $S$ has dense range, so the range projection of $S$ is $I_n$), we get that

$$\begin{bmatrix} 0 & 1_{[0,1-\delta]}(C)\,C \\ 0 & 0 \end{bmatrix} \in M_2(W^*(C)).$$

Making $\delta \to 0$ and then using a similar trick for $C$ as we just did for $S$, it follows (since $C$ has dense range) that

$$\begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \in M_2(W^*(C)).$$

This implies that the four operator-block matrix units $E_{11}, E_{12}, E_{21}, E_{22}$ are in $M_2(W^*(C))$. Now given $X = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \in W^*(P_0, Q_0)$, we know that $A_1 \in P_0\mathcal{M}P_0$. Also

$$\begin{bmatrix} A_2 & 0 \\ 0 & 0 \end{bmatrix} = E_{11}XE_{21} \in M_2(W^*(C)).$$

So $A_2 \in W^*(C)$, and similarly $A_3, A_4 \in W^*(C)$. Hence $Z\mathcal{M} \subset M_2(W^*(C))$, proving they are equal. □

**14.3.2. Type II von Neumann Algebras.** By definition, a type II von Neumann algebra contains, below any fixed projection $P$, a strictly decreasing sequence of projections $P \geq P_1 \geq P_2 \geq \cdots$. It is easy to find chains like this in $\mathcal{B}(\mathcal{H})$. For example, in terms of a fixed family of matrix units, let $P_n \in \mathcal{B}(\ell^2(\mathbb{N}))$ be

$$P_n = \sum_{k=1}^{\infty} P_{2^n k, 2^n k}.$$

This is a monotone decreasing sequence of projections. Of course they are all infinite, and $P_n \sim P_m$ for all $n, m$. In a type $\mathrm{II}_1$ algebra, though, all these projections would be finite and hence no two of them would be equivalent, even if they are equivalent in $\mathcal{B}(\mathcal{H})$.

Besides its intrinsic value, one immediate consequence of Proposition 14.2.22 is

*14.3.7. Corollary.*

*If $\mathcal{M}$ has no type I component, then it is diffuse.*

The converse of Corollary 14.3.7 does not hold; for instance $L^\infty[0,1]$ is type I and diffuse.

*14.3.8. Remark .* Consider the von Neumann algebra $L(G)$, for $G$ an icc group, as in Example 13.3.13. We know this is finite, infinite dimensional, and a factor. By the Classification Theorem, we know that $L(G)$ is of a fixed type. It cannot be $\mathrm{I}_n$ because it is infinite-dimensional. It cannot be $\mathrm{I}_\infty$, because it would not be finite. So the only possibility is that it is $\mathrm{II}_1$. It seems like we then have many examples of $\mathrm{II}_1$ factors; and we do, but this is a tricky business. Murray and von Neumann [MvN43] were able to prove $L(S_\infty) \not\simeq L(\mathbb{F}_2)$, where $S_\infty$ is the group of finite permutations of $\mathbb{N}$. They were also able to prove that there is a unique approximately finite-dimensional $\mathrm{II}_1$-factor (that is, with a wot-dense increasing family of finite-dimensional subalgebras), with $L(S_\infty)$ being one realization of it; this is known as the **hyperfinite $\mathrm{II}_1$-factor**. For many years it was really hard to come up with examples. To this day, it is an open question whether $L(\mathbb{F}_n) \simeq L(\mathbb{F}_m)$ if $m \neq n$. That said, there have been spectacular advances in our understanding of $\mathrm{II}_1$-factors. In the late 1960s McDuff found a countable family of pairwise non-isomorphic $\mathrm{II}_1$-factors [McD69], soon extended to uncountable—using the same ideas—by Sakai [Sak70]. By then it was known that $L(G)$ is the hyperfinite $\mathrm{II}_1$-factor whenever $G$ is amenable and icc. In 1976 Connes [Con76], among many brilliant results, proved the converse, so $L(G)$ is hyperfinite if and only if $G$ is amenable. Since then, the study of $\mathrm{II}_1$-factors grew first into Jones' Index Theory (started in [Jon83]) and Popa's Deformation/Rigidity (started in [Pop06a, Pop06b, Pop07]), and both areas keep thriving. A third area, Free Probability, grew out of Voiculescu's efforts to address the isomorphism question for free group factors.

If we take a $\mathrm{II}_1$ von Neumann algebra $\mathcal{M}$ and $\mathcal{H}$ with $\dim\mathcal{H} = \infty$, then $\mathcal{M} \otimes \mathcal{B}(\mathcal{H})$ is type $\mathrm{II}_\infty$. It turns out that this is the only way a $\mathrm{II}_\infty$ algebra can be, up to direct sum/direct integrals (we will prove the factor case below). And, way beyond the scope of this text (as we are ignoring Tomita–Takesaki Modular Theory), a type III factor can be written as a crossed product of a $\mathrm{II}_\infty$ factor by $\mathbb{R}$; so in a sense $\mathrm{II}_1$ factors are the building blocks of every non-type I von Neumann algebra.

***14.3.9. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a $II_\infty$ factor. Then there exists a $II_1$-factor $\mathcal{N}$ and a Hilbert space $\mathcal{K}$ such that $\mathcal{M} \simeq \mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$.*

*Proof.* By definition there exists a finite projection $P_0$ in $\mathcal{M}$. Using Proposition 14.2.23 we get a family $\{P_j\}_{j\in J} \subset \mathcal{M}$ of pairwise orthogonal projections, with $\sum_j P_j = I_\mathcal{M}$ and with $P_j \sim P_0$ for all $j$. Let $j_0 \in J$ be the index such that $P_{j_0} = P_0$. We have partial isometries $\{E_{j_0,j}\}_{j\in J}$ with $E_{j_0,j}^* E_{j_0,j} = P_j$ and $E_{j_0,j}E_{j_0,j}^* = P_0$. This allows us to define matrix units by

$$E_{kj} = E_{j_0,k}^* E_{j_0,j}.$$

Let $\mathcal{N} = P_0\mathcal{M}P_0$. We know from Corollary 14.1.4 that $\mathcal{N}$ is a factor. By Proposition 14.3.4, it is type II. And since $P_0$ is finite, it is $\mathrm{II}_1$. Now using Exercise 13.4.5 we have that $\mathcal{M} \simeq \mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\ell^2(J))$. $\square$

When $\mathcal{M}$ is not a factor in Proposition 14.3.9, the result becomes instead that $\mathcal{M}$ is of the form $\prod_\alpha Z_\alpha \mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H}_\alpha)$, where $\mathcal{N}$ is a type $\mathrm{II}_1$ von Neumann algebra, $\{Z_\alpha\}$ are pairwise orthogonal central projections in $\mathcal{N}$ that add to $I_\mathcal{N}$, and $\alpha$ are cardinals with $\dim \mathcal{H}_\alpha = \alpha$.

Neither the choice of $\mathcal{N}$ nor $\mathcal{K}$ is canonical in Proposition 14.3.9. It is not hard to check that if $\dim \mathcal{K} = \infty$ then $\mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}) \simeq M_2(\mathcal{N}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$, for instance (Exercise 14.3.7). And there are known examples of $\mathrm{II}_1$-factors $\mathcal{N}$ such that $\mathcal{N} \not\simeq M_2(\mathcal{N})$ (one such example would be a $\mathrm{II}_1$-factor with trivial **fundamental group** (see subsection 14.7.1); the first such factor, determined in [Pop06a], is $L(\mathbb{Z}_2 \rtimes SL_2(\mathbb{Z})))$.

We know from Exercise 13.3.8 that the group $\mathrm{II}_1$-factors considered in Example 13.3.13 have a faithful tracial state. It turns out that this is always the case for $\mathrm{II}_1$-factors. This will follow from the existence of the centre-valued trace in finite von Neumann algebras (Theorem 14.5.10). Restricted to the set of projections, a trace gives a **dimension function**. The same way that the trace in $\mathcal{B}(\mathcal{H})$ quantifies the size of a projection, a similar function can be defined on any semifinite von Neumann algebra. In a $\mathrm{II}_1$-factor this dimension function has continuous range, namely the whole interval $[0, 1]$. We outline an ad-hoc construction of the dimension function for $\mathrm{II}_1$-factors in Exercise 14.3.6.

**14.3.3. The Type of the Commutant.** We now focus on the unexpected result that the commutant of a von Neumann algebra $\mathcal{M}$ is always of the same major type (I, II, III) than $\mathcal{M}$.

**14.3.10. Lemma.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra with cyclic and separating vector $\xi \in \mathcal{H}$. Suppose that $\mathcal{M}'$ is properly infinite. Then no nonzero projection $Z \in \mathcal{P}(\mathcal{Z}(\mathcal{M}))$ is abelian on $\mathcal{M}$.*

*Proof.* Suppose that $Z \in \mathcal{Z}(\mathcal{M})$ is a nonzero projection with $\mathcal{M}Z$ abelian. As $Z\xi$ is separating for $\mathcal{M}Z$, by Exercise 12.5.5 we know that $\mathcal{M}Z$ is $\sigma$-finite. As $Z\xi$ is cyclic for $\mathcal{M}Z$ in $\mathcal{B}(Z\mathcal{H})$, by Proposition 12.5.10 we get that $(\mathcal{M}Z)' = \mathcal{M}Z$, so $\mathcal{M}Z = (\mathcal{M}Z)' = \mathcal{M}'Z$ (the last equality by Proposition 14.1.1). This means that $\mathcal{M}'Z$ is abelian, contradicting—since $Z$ is central—the assumption that $\mathcal{M}'$ is properly infinite. □

**14.3.11. Lemma.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a finite von Neumann algebra and $\xi \in \mathcal{H}$ cyclic and separating for $\mathcal{M}$. Then $\mathcal{M}'$ is finite.*

*Proof.* We know from Proposition 12.5.2 that $\xi$ is cyclic and separating for $\mathcal{M}'$. Suppose that $\mathcal{M}'$ is not finite. By Proposition 14.2.13 there exists a central projection $Z \in \mathcal{Z}(\mathcal{M}')$ with $\mathcal{M}'Z$ properly infinite. We have, using that $\xi$ is cyclic for $\mathcal{M}'$,

$$\overline{\mathcal{M}'Z\xi} = \overline{Z\mathcal{M}'\xi} = Z\overline{\mathcal{M}'\xi} = Z\mathcal{H}.$$

We similarly get that $\xi$ is cyclic for $\mathcal{M}Z$. By Proposition 12.5.2 it is also separating for $\mathcal{M}Z$ and $\mathcal{M}'Z$. If we show that $\mathcal{M}Z$ is not finite, then $\mathcal{M}$ will not be finite. Hence by working with $\mathcal{M}Z$ and $\mathcal{M}'Z$ we may assume without loss of generality that $\mathcal{M}'$ is properly infinite and show that this implies that $\mathcal{M}$ is not finite.

Now we use Zorn's Lemma to construct a maximal family of orthogonal projections $\{P_j\} \subset \mathcal{M}$ with pairwise orthogonal central supports and such that each $P_j$ is a proper subprojection of $c(P_j)$ for all $j$ (the family necessarily exists for otherwise there would exist a central projection $Z \in \mathcal{Z}(\mathcal{M})$ with $\mathcal{M}Z$ abelian, contradicting Lemma 14.3.10). Let $P = \sum_j P_j \in \mathcal{M}$. Since $c(P) = \sum_j c(P_j)$, we have that $P$ is a proper subprojection of $c(P)$. Any subprojection $Z \leq I_{\mathcal{M}} - c(P)$ will satisfy, by the maximality, $Z = c(Z) \in \mathcal{Z}(\mathcal{M})$. So every projection in $\mathcal{M}(I_{\mathcal{M}} - c(P))$ is abelian, which makes the whole algebra $\mathcal{M}(I_{\mathcal{M}} - c(P))$ abelian. By Lemma 14.3.10, $I_{\mathcal{M}} - c(P) = 0$; that is, $c(P) = I_{\mathcal{M}}$.

Let $Q = [\mathcal{M}P\xi] \in \mathcal{P}(\mathcal{M}')$. Because $\xi$ is separating for $\mathcal{M}$, the algebra is $\sigma$-finite by Exercise 12.5.5. Then Corollary 14.2.18 gives us that $Q \sim c(Q)$.

Hence, using that $\xi$ is cyclic for $\mathcal{M}'$,

$$Q \sim c(Q) = [\mathcal{M}'\mathcal{M}P\xi] = [\mathcal{M}P\mathcal{M}'\xi] = [\mathcal{M}P\mathcal{H}] = c(P) = I_{\mathcal{M}}.$$

That is $[\mathcal{M}P\xi] \sim I_{\mathcal{M}} = [\mathcal{M}\xi]$ in $\mathcal{M}'$. By Lemma 14.2.25,

$$P = [P\mathcal{M}'\xi] = [\mathcal{M}'P\xi] \sim [\mathcal{M}'\xi] = I_{\mathcal{M}}$$

in $\mathcal{M}$. As $P$ is a proper subprojection of $c(P) = I_{\mathcal{M}}$, this implies that $\mathcal{M}$ is not finite. □

**_14.3.12. Proposition._**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate von Neumann algebra, $\xi \in \mathcal{H}$. The projection $[\mathcal{M}\xi]$ is abelian, finite, infinite, or properly infinite in $\mathcal{M}'$, if and only if $[\mathcal{M}'\xi]$ has the same property in $\mathcal{M}$.*

*Proof.* Write $P = [\mathcal{M}'\xi] \in \mathcal{M}$, $Q = [\mathcal{M}\xi] \in \mathcal{M}'$. Since

$$c(P) = [\mathcal{M}\mathcal{M}'\xi] = [\mathcal{M}'\mathcal{M}\xi] = c(Q),$$

by working on $\mathcal{M}c(P)$ and $\mathcal{M}'c(P)$ we may assume without loss of generality that $c(P) = c(Q) = I_{\mathcal{H}}$. The isomorphism $\gamma : \mathcal{M}' \to \mathcal{M}'P$ given by $\gamma(S) = SP$ (Corollary 14.1.4) has $\gamma(Q) = QP$. Hence $Q$ is abelian or finite or infinite or properly infinite in $\mathcal{M}'$ if and only if the same holds for $QP$ in $\mathcal{M}'P$. As $P\xi = \xi$, using Exercise 14.2.11

$$[\mathcal{M}'P\xi] = [\mathcal{M}'\xi] = P, \qquad [P\mathcal{M}P\xi] = [P\mathcal{M}\xi] = P[\mathcal{M}\xi] = PQ = QP.$$

All the properties we are considering (abelian, finite, infinite, properly infinite) are defined in terms of what happens below $P$; that is, in $P\mathcal{M}P$. Similarly for $QP$, we can work on $\mathcal{M}'P$. As $\mathcal{M}'P = (P\mathcal{M}P)'$, we are reducing to the case where $P = I_{\mathcal{M}}$, and we are trying to show that $Q \in \mathcal{M}'$ with $c(Q) = I_{\mathcal{M}}$ is abelian, finite, infinite, or properly infinite in $\mathcal{M}'$ if and only if $I_{\mathcal{M}}$ is in $\mathcal{M}$. If we then we rehash the argument with the roles of $I_{\mathcal{M}}$ and $Q$ reversed, we conclude that we can assume that $P = Q = I_{\mathcal{M}}$. That is, we want to show that $\mathcal{M}$ is abelian, finite, infinite, or properly infinite if and only if $\mathcal{M}'$ is, in the presence of a cyclic vector $\xi$. As $\xi$ is cyclic for both $\mathcal{M}$ and $\mathcal{M}'$, it is also separating for both (Proposition 12.5.2).

If $\mathcal{M}$ is abelian, the presence of a separating vector guarantees that it is $\sigma$-finite (Exercise 12.5.5). And then the presence of a cyclic vector forces $\mathcal{M}'$ is abelian by Proposition 12.5.10.

If $\mathcal{M}$ is finite, then $\mathcal{M}'$ is finite by Lemma 14.3.11. And similarly if $\mathcal{M}'$ is finite, then $\mathcal{M}$ is finite. This immediately implies that $\mathcal{M}$ is infinite if and only if $\mathcal{M}'$ is infinite. And if $\mathcal{M}$ is properly infinite and $Q \in \mathcal{M}'$ is a central projection, if $\mathcal{M}'Q$ were finite then so would $\mathcal{M}Q = (\mathcal{M}'Q)'$ be, contradicting that $\mathcal{M}$ is properly infinite. It follows that $\mathcal{M}$ is properly infinite if and only if $\mathcal{M}'$ is. □

***14.3.13. Theorem.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a non-degenerate von Neumann algebra. Then $\mathcal{M}$ is of one of the types I, II, III if and only if $\mathcal{M}'$ is of the same type.*

*Proof.* Suppose that $\mathcal{M}$ is type I. Then there exists an abelian projection $P \in \mathcal{M}$ with $c(P) = I_{\mathcal{M}}$ (Exercise 14.3.2). If $\mathcal{M}'$ is not type I, by Theorem 14.3.2 there exists a central projection $Q \in \mathcal{M}'$ such that $\mathcal{M}'Q$ is type II or type III. The fact that $c(P) = I_{\mathcal{M}}$ guarantees that $QP \neq 0$. Fix nonzero $\xi \in QP\mathcal{H}$. Consider the projections $[\mathcal{M}'\xi] \in \mathcal{M}Q$ and $[\mathcal{M}\xi] \in \mathcal{M}'Q = (\mathcal{M}Q)'$. As $[\mathcal{M}'\xi] = [\mathcal{M}'P\xi] = P[\mathcal{M}'\xi] \leq P$ (using Exercise 14.2.11), the projection $[\mathcal{M}'\xi]$ is abelian in $\mathcal{M}Q$. By Proposition 14.3.12, $[\mathcal{M}\xi]$ is abelian in $\mathcal{M}'Q$, contradicting that $\mathcal{M}'Q$ is type II or III. Thus $\mathcal{M}'$ is type I.

If $\mathcal{M}$ is type II, the argument runs very similar. There exists a finite projection $P \in \mathcal{M}$ with $c(P) = I_{\mathcal{M}}$ (Exercise 14.3.5). If $\mathcal{M}'$ is not type II, by Theorem 14.3.2 there exists a nonzero central projection $Q \in \mathcal{M}'$ such that $\mathcal{M}'Q$ is type I or type III; but it cannot be type I because then $\mathcal{M}Q$ would be type I by the first paragraph. The fact that $c(P) = I_{\mathcal{M}}$ guarantees that $QP \neq 0$. Fix nonzero $\xi \in QP\mathcal{H}$. Consider the projections $[\mathcal{M}'\xi] \in \mathcal{M}Q$ and $[\mathcal{M}\xi] \in \mathcal{M}'Q = (\mathcal{M}Q)'$. As $[\mathcal{M}'\xi] = [\mathcal{M}'P\xi] = P[\mathcal{M}'\xi] \leq P$, the projection $[\mathcal{M}'\xi]$ is finite in $\mathcal{M}Q$ (Exercise 14.2.17). By Proposition 14.3.12, $[\mathcal{M}\xi]$ is finite in $\mathcal{M}'Q$, contradicting that $\mathcal{M}'Q$ is type III. Thus $\mathcal{M}'$ is type II.

Finally, if $\mathcal{M}$ is type III then $\mathcal{M}'$ is necessarily is type III, for otherwise $\mathcal{M}$ would not be type III by the previous parts of the proof. □

The fact that the commutant of a von Neumann algebra is of the same type cannot be extended to the subtypes $\mathrm{I}_n$ and $\mathrm{II}_1$ and $\mathrm{II}_\infty$. For instance if $\mathcal{M} = \mathcal{B}(\mathcal{H})$ is $I_\infty$ then $\mathcal{M}' = \mathbb{C}\, I_{\mathcal{H}}$ of type $\mathrm{I}_1$. In fact, using Exercise 13.4.27, given $\mathcal{M}$ of type $\mathrm{I}_m$ and $\mathcal{N}$ of type $\mathrm{I}_n$ (infinity allowed), we have $\mathcal{M} \otimes I_{\mathcal{N}}$ of type $\mathrm{I}_m$ with commutant $I_{\mathcal{M}} \otimes \mathcal{N}$ of type $\mathrm{I}_n$. If $\mathcal{M}$ is a $\mathrm{II}_1$ von Neumann algebra we can consider $\mathcal{M} \otimes I_{\mathcal{H}} \subset \mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$; this is a $\mathrm{II}_1$ von Neumann algebra with its commutant containing $I_{\mathcal{M}} \otimes \mathcal{B}(\mathcal{H})$ so it is $\mathrm{II}_\infty$.

**14.3.4. Further Classification Refinements.** Murray and von Neumann produced examples of all types of von Neumann algebras already in the 1930s. But type III remained entirely elusive for more than 30 years afterwards (see Section 14.6). It was only in the late 1960s, with Tomita's unpublished work that was collected and expanded by Takesaki [Tak70] that things began to unravel. The original impact of Tomita's work was a proof that the commutant of the tensor product is the tensor product of the commutants; and also the fact that if $\mathcal{M}$ is represented in a way so that it has a cyclic and separating vector, then $\mathcal{M}'$ is anti-isomorphic to $\mathcal{M}$. The techniques were far reaching though, and soon all kinds of new results were produced. Powers [Pow67] had found uncountably many non-isomorphism type III factors (constructed as infinite tensor products of type I factors), and Araki–Woods [AW69] produced a classification of such factors. Most notably Connes [Con71, Con73], building of work of Arveson [Arv74], produced a classification of factors of type III into type $\mathrm{III}_\lambda$, $\lambda \in [0,1]$. All of [Tak03] is devoted to these topics.

**14.3.5. Exercises.**

(14.3.1) Show that the matrix units defined in (14.5) do satisfy the matrix unit relations $E_{kj}E_{ab} = \delta_{j,a}E_{kb}$ and $E_{kj}^* = E_{jk}$.

(14.3.2) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that $\mathcal{M}$ is type I if and only if there exists a projection $P \in \mathcal{M}$, abelian, with $c(P) = I_{\mathcal{M}}$.

(14.3.3) In the proof of Proposition 14.3.6, show that $P_0$ and $Q_0$ are in generic position when acting on $(P_0 \vee Q_0)\mathcal{H}$.

(14.3.4) Let $\mathcal{M}$ be a type I von Neumann algebra and $\mathcal{H}$ a Hilbert space. Show that $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$ is type I, without using Theorem 14.3.2.

(14.3.5) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Show that $\mathcal{M}$ is type II if and only if there exists a projection $P \in \mathcal{M}$, finite, with $c(P) = I_{\mathcal{M}}$.

(14.3.6) Let $\mathcal{M}$ be a $\mathrm{II}_1$-factor. We will outline here a way to "manually" construct the normalized dimension function.

(i) Use Proposition 14.2.22 to construct a family $\{P_{k,n}\} \subset \mathcal{M}$ of projections, with $n \in \mathbb{N}$, $k \in \{1, \ldots, 2^n\}$, and

$$P_{k,n} \sim P_{j,n}, \qquad P_{2k-1,n+1} + P_{2k,n} = P_{k,n}$$

for all $n$ and all $k, j \leq 2^n$, and $\sum_k P_{k,n} = I_{\mathcal{M}}$ (note: $\{P_{k,n}\}_{k=1}^{2^n}$ are pairwise orthogonal by Proposition 10.5.5).

(ii) (Division Algorithm) Show that given $n \in \mathbb{N}$ and $P \in \mathcal{P}(\mathcal{M})$, there exist $s(n) \in \{0, \ldots, 2^n\}$ and projections

$$Q_1, \ldots, Q_{s(n)}, R \in \mathcal{M},$$

with $R \prec P_{1,n}$, $Q_k \sim P_{1,n}$ for all $k$, and such that $P = R + \sum_{k \leq s(n)} Q_k$.

(iii) Show that if $Q_1', \ldots, Q_{s'(n)}', R'$ is another decomposition for $P$ as above, then $s'(n) = s(n)$ and $R' \sim R$.

(iv) Keep considering the same fixed $P$. Show that $s(n+1) \geq 2s(n)$ for all $n$.

(v) For the same projection $P$, let $\alpha_n = 2^{-n}\, s(n)$. Show that the sequence $\{\alpha_n\}$ converges to some number $\tau(P) \in [0,1]$.

(vi) Show that for projections $P, Q \in \mathcal{M}$, we have $P \preceq Q$ if and only if $\tau(P) \leq \tau(Q)$. Conclude that $P \sim Q$ if and only if $\tau(P) = \tau(Q)$ and that $P \prec Q$ if and only if $\tau(P) < \tau(Q)$.

(vii) Use Proposition 14.2.16 to show that if $\{Q_n\} \subset \mathcal{M}$ is a monotone sequence of projections with $Q_n \xrightarrow{\text{sot}} Q$, then $\tau(Q_n) \to \tau(Q)$.

(viii) Show that $\tau$ is $\sigma$-additive.

(ix) Show that $\tau(\mathcal{P}(\mathcal{M})) = [0,1]$.

(14.3.7) Let $\mathcal{N}$ be a von Neumann algebra and $\mathcal{K}$ an infinite-dimensional Hilbert space. Show that $\mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K}) \simeq M_2(\mathcal{N}) \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$.

## 14.4. Tensor Products of von Neumann Algebras

This section should be considered as a reference; it is not really needed at this stage if the goal is to learn about von Neumann algebras. Just knowing that Theorem 14.4.3 exists should be enough.

We defined the tensor product $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N}$ of von Neumann algebras in Definition 13.4.6. Namely, if $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ then $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N} \subset \mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$ is defined to be $(\mathcal{M} \otimes \mathcal{N})''$.

As an exception to our rule, in order to prove Theorem 14.4.3, it makes sense for our current purposes to consider real Hilbert spaces; we saw those briefly in Exercise 4.3.20. Given a real subspace $K \subset \mathcal{H}$, its **real orthogonal** is

$$K_{\mathbb{R}}^{\perp} = \{\eta \in \mathcal{H} : \ \eta \perp_{\mathbb{R}} K\} = \{\eta \in \mathcal{H} : \ \operatorname{Re}\langle \eta, \xi\rangle = 0, \ \xi \in K\}.$$

The reason we care about this is that if $S, T \in \mathcal{B}(\mathcal{H})$ are selfadjoint then they commute if and only if $S\xi \perp_{\mathbb{R}} iT\xi$ for all $\xi \in \mathcal{H}$ (Exercise 14.4.1).

Most of the technicalities we need are contained in the next two lemmas.

***14.4.1. Lemma.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra with cyclic vector $\xi_0$, and $\mathcal{A} \subset \mathcal{M}$, $\mathcal{B} \subset \mathcal{M}'$ non-degenerate $*$-subalgebras. The following statements are equivalent:*

*(i)* $\mathcal{A}'' = \mathcal{M}$ *and* $\mathcal{B}'' = \mathcal{M}'$;

*(ii)* $\overline{\mathcal{A}^{\rm sa}\xi_0 + i\mathcal{B}^{\rm sa}\xi_0} = \mathcal{H}$;

*(iii)* $(\mathcal{A}^{\rm sa}\xi_0)_{\mathbb{R}}^{\perp} = \overline{i\mathcal{B}^{\rm sa}\xi_0}$.

*Proof.* (i) $\Longrightarrow$ (ii) By the Double Commutant Theorem (12.3.5) we know that $\overline{\mathcal{A}}^{\rm sot} = \mathcal{M}$ and $\overline{\mathcal{B}}^{\rm sot} = \mathcal{M}'$. Combined with the fact that $\xi_0$ is cyclic for $\mathcal{M}$

so $\overline{\mathcal{M}\xi_0} = \mathcal{H}$, we get that $\overline{\mathcal{A}\xi_0} = \mathcal{H}$. We know from Exercise 14.4.1 that $\mathcal{A}^{\rm sa}\xi_0 \perp_{\mathbb{R}} i\mathcal{B}^{\rm sa}\xi_0$. Suppose that $\eta \in \mathcal{H}$ is real orthogonal to $\mathcal{A}^{\rm sa}\xi_0 + i\mathcal{B}^{\rm sa}\xi_0$. By Kaplansky (Theorem 12.3.17) we have that $\mathcal{M}^{\rm sa} = \overline{(\mathcal{A}^{\rm sa})}^{\rm sot}$ and hence $\eta$ is real orthogonal to $\mathcal{M}_{\rm sa}\xi_0 + i\mathcal{M}'_{\rm sa}\xi_0$ (we switch for now to subscripts for indicating the selfadjoint part due to the conflict with the prime in $\mathcal{M}'$). Let $\tilde\pi : \mathcal{M} \to \mathcal{B}(\mathcal{H} \otimes \mathbb{C}^2)$ be the 2-amplification of the identity representation of $\mathcal{M}$; that is, $\tilde\pi = \pi \otimes I_2$. We know from Exercise 13.4.27 that $\tilde\pi(\mathcal{M})' = \mathcal{M}' \otimes M_2(\mathbb{C}) = M_2(\mathcal{M}')$. Let $\zeta_0 = \xi_0 \otimes e_1 + \eta \otimes e_2$ and let $P$ be the orthogonal projection onto $\overline{\tilde\pi(\mathcal{M})\zeta_0}$. By Exercise 11.6.8 we know that $P \in \tilde\pi(\mathcal{M})'$. This allows us to write

$$P = \begin{bmatrix} X & Y \\ Y^* & Z \end{bmatrix}$$

for some $X, Y, Z \in \mathcal{M}'$, with $0 \leq X, Z \leq I_{\mathcal{H}}$. We have by construction that $P\zeta_0 = \zeta_0$. This gives the two equalities

$$X\xi_0 + Y\eta = \xi_0 \qquad \text{and} \qquad Y^*\xi_0 + Z\eta = \eta. \tag{14.9}$$

Our hypothesis is that $\eta \perp_{\mathbb{R}} \mathcal{M}_{\rm sa}\xi_0$ and $\eta \perp_{\mathbb{R}} \mathcal{M}'_{\rm sa}\xi_0$. The first orthogonality says that for each $T \in \mathcal{M}_{\rm sa}$ we have $\operatorname{Re}\langle T\xi_0, \eta\rangle = 0$; this means that $\langle T\xi_0, \eta\rangle$ is imaginary, and then—using also that $T = T^*$—

$$\langle T\xi_0, \eta\rangle = -\overline{\langle T\xi_0, \eta\rangle} = -\langle T\eta, \xi_0\rangle.$$

The equality between the first and last term is linear on $T$, so it holds for all $T \in \mathcal{M}$. Hence

$$\begin{aligned} \langle \tilde\pi(T)\zeta_0, \eta \otimes e_1 + \xi_0 \otimes e_2\rangle &= \langle T\xi_0 \otimes e_1 + T\eta \otimes e_2, \eta \otimes e_1 + \xi_0 \otimes e_2\rangle \\ &= \langle T\xi_0, \eta\rangle + \langle T\eta, \xi_0\rangle = 0 \end{aligned}$$

for all $T \in \mathcal{M}$. This shows that $\eta \otimes e_1 + \xi_0 \otimes e_2$ is orthogonal to the range of $P$, so $P(\eta \otimes e_1 + \xi_0 \otimes e_2) = 0$. Then

$$X\eta + Y\xi_0 = 0. \tag{14.10}$$

We also have that $\eta$ is real orthogonal to $i\mathcal{M}'_{\rm sa}\xi_0$. So for every $S \in \mathcal{M}'_{\rm sa}$ we have $0 = \operatorname{Re}\langle \eta, iS\xi_0\rangle = \operatorname{Im}\langle \eta, S\xi_0\rangle$. That is, $\langle \eta, S\xi_0\rangle$ is real and so

$$\langle S\eta, \xi_0\rangle = \langle \xi_0, S\eta\rangle = \langle S\xi_0, \eta\rangle. \tag{14.11}$$

Since $X, Y, Z \in \mathcal{M}'$, (14.11) applies to each of them. Thus, using first (14.10), then (14.11), and then (14.9),

$$\begin{aligned} 0 \leq \langle X\eta, \eta\rangle &= -\langle Y\xi_0, \eta\rangle = -\langle Y\eta, \xi_0\rangle \\ &= -\langle (I_{\mathcal{H}} - X)\xi_0, \xi_0\rangle \leq 0. \end{aligned}$$

We get that $X\eta = 0$ and that $(I_{\mathcal{H}} - X)\xi_0 = 0$. As $\xi_0$ is separating for $\mathcal{M}'$ (Proposition 12.5.2), this implies that $X = I_{\mathcal{H}}$ and therefore $\eta = 0$.

(ii) $\implies$ (iii) As already mentioned, from Exercise 14.4.1 we know that $i\mathcal{B}^{\mathrm{sa}}\xi_0 \subset (\mathcal{A}^{\mathrm{sa}}\xi_0)^{\perp}_{\mathbb{R}}$. Given $\xi \in \mathcal{H}$, we have by hypothesis $\xi = \lim_n T_n\xi_0 + iS_n\xi_0$ with $\{T_n\} \subset \mathcal{A}^{\mathrm{sa}}$ and $\{S_n\} \subset \mathcal{B}^{\mathrm{sa}}$. Since for any $T_0 \in \mathcal{A}$ and $S_0 \in \mathcal{B}$

$$\begin{aligned}\|T_0\xi_0 + iS_0\xi_0\|^2 &= \|T_0\xi_0\|^2 + \|S_0\xi_0\|^2 + 2\mathrm{Re}\,\langle T_0\xi_0, iS_0\xi_0\rangle \\ &= \|T_0\xi_0\|^2 + \|S_0\xi_0\|^2,\end{aligned}$$

it follows that $\{T_n\xi_0\}$ and $\{S_n\xi_0\}$ are both Cauchy. So $\xi = \xi_1 + \xi_2$, with $\xi_1 \in \overline{\mathcal{A}^{\mathrm{sa}}\xi_0}$ and $\xi_2 \in \overline{\mathcal{B}^{\mathrm{sa}}\xi_0}$. That is,

$$\mathcal{H} = \overline{\mathcal{A}^{\mathrm{sa}}\xi_0} + \overline{i\mathcal{B}^{\mathrm{sa}}\xi_0}.$$

If now $\eta \in (\mathcal{A}^{\mathrm{sa}}\xi_0)^{\perp}_{\mathbb{R}}$ we can write $\eta = \eta_1 + \eta_2$, with $\eta_1 \in \overline{\mathcal{A}^{\mathrm{sa}}\xi_0}$ and $\eta_2 \in \overline{i\mathcal{B}\xi_0}$. But then $\mathrm{Re}\,\langle \eta, \eta_1\rangle = 0$. So

$$\begin{aligned}\|\eta\|^2 = \langle \eta, \eta\rangle &= \mathrm{Re}\,\langle \eta, \eta_1\rangle + \mathrm{Re}\,\langle \eta, \eta_2\rangle \\ &= \mathrm{Re}\,\langle \eta, \eta_2\rangle = \mathrm{Re}\,\langle \eta_2, \eta_2\rangle = \|\eta_2\|^2.\end{aligned}$$

Thus $\eta_1 = 0$ and $\eta = \eta_2 \in \overline{i\mathcal{B}^{\mathrm{sa}}\xi_0}$.

(iii) $\implies$ (i) We have by Exercise 14.4.1

$$i\mathcal{A}'_{\mathrm{sa}}\xi_0 \subset (\mathcal{A}^{\mathrm{sa}}\xi_0)^{\perp}_{\mathbb{R}} = \overline{i\mathcal{B}^{\mathrm{sa}}\xi_0}.$$

We deduce that $\mathcal{A}'_{\mathrm{sa}}\xi_0 \subset \overline{\mathcal{B}^{\mathrm{sa}}\xi_0}$. Hence given $T \in \mathcal{A}'_{\mathrm{sa}}$ there exists a sequence $\{S_n\} \subset \mathcal{B}^{\mathrm{sa}}$ such that $T\xi_0 = \lim_n S_n\xi_0$. Given $R \in \mathcal{B}'$ and $X, Y \in \mathcal{A}$, and using that $\mathcal{B} \subset \mathcal{A}'$,

$$\begin{aligned}\langle RTX\xi_0, Y\xi_0\rangle &= \langle RXT\xi_0, Y\xi_0\rangle = \lim_n \langle RXS_n\xi_0, Y\xi_0\rangle \\ &= \lim_n \langle RX\xi_0, YS_n\xi_0\rangle = \langle RX\xi_0, YT\xi_0\rangle \qquad (14.12) \\ &= \langle TRX\xi_0, Y\xi_0\rangle.\end{aligned}$$

Also,

$$\mathcal{M}_{\mathrm{sa}}\xi_0 \subset (i\mathcal{M}'_{\mathrm{sa}}\xi_0)^{\perp}_{\mathbb{R}} \subset (i\mathcal{B}_{\mathrm{sa}}\xi_0)^{\perp}_{\mathbb{R}} = (\mathcal{A}^{\mathrm{sa}}\xi_0)^{\perp\perp}_{\mathbb{R}} = \overline{\mathcal{A}_{\mathrm{sa}}\xi_0}.$$

This implies

$$\mathcal{M}\xi_0 = \mathcal{M}_{\mathrm{sa}}\xi_0 + i\mathcal{M}_{\mathrm{sa}}\xi_0 \subset \overline{\mathcal{A}^{\mathrm{sa}}\xi_0} + i\overline{\mathcal{A}^{\mathrm{sa}}\xi_0} \subset \overline{\mathcal{A}\xi_0},$$

which shows that $\xi_0$ is cyclic for $\mathcal{A}$. If we now go back to (14.12) this means that $TR = RT$ for all $T \in \mathcal{A}'_{\mathrm{sa}}$ (and a fortiori for all $T \in \mathcal{A}'$) and $R \in \mathcal{B}'$. From this we get

$$\mathcal{A}' \subset (\mathcal{B}')' = \mathcal{B}'' \subset \mathcal{M}' \subset \mathcal{A}', \qquad \mathcal{B}' \subset (\mathcal{A}')' = \mathcal{A}'' \subset \mathcal{M} \subset \mathcal{B}'.$$

Therefore $\mathcal{A}'' = \mathcal{M}'' = \mathcal{M}$ and $\mathcal{B}'' = \mathcal{M}'$. ∎

*14.4.2. Lemma.*

*Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces and $H\subset\mathcal{H}$ and $K\subset\mathcal{K}$ real subspaces such that $H+iH$ and $K+iK$ are respectively dense in $\mathcal{H}$ and $\mathcal{K}$. Then*

$$\overline{H\otimes K+i(H_{\mathbb{R}}^{\perp}\otimes K_{\mathbb{R}}^{\perp})}=\mathcal{H}\mathbin{\overline{\otimes}}\mathcal{K}.$$

*Proof.* We want to show that the orthogonal of $H\otimes K+i(H_{\mathbb{R}}^{\perp}\otimes K_{\mathbb{R}}^{\perp})$ is $\{0\}$. So let $\zeta\in\mathcal{H}\overline{\otimes}\mathcal{K}$ with $\zeta\perp_{\mathbb{R}} H\otimes K$ and $\zeta\perp_{\mathbb{R}} i(H_{\mathbb{R}}^{\perp}\otimes K_{\mathbb{R}}^{\perp})$. Define a (real) bilinear form on $\mathcal{H}\times\mathcal{K}$ by $B(\xi,\eta)=\operatorname{Re}\langle\zeta,\xi\otimes\eta\rangle$. This is bounded, for $|B(\xi,\eta)|\leq\|\zeta\|\,\|\xi\|\,\|\eta\|$. The proof of the Riesz Representation Theorem (4.5.4) does not really use that the inner product is sesquilinear, so it works without change in a real Hilbert space; the same happens with Proposition 10.1.5. So there exists a real linear bounded operator $T:\mathcal{H}\to\mathcal{K}$ such that $\operatorname{Re}\langle\zeta,\xi\otimes\eta\rangle=\operatorname{Re}\langle T\xi,\eta\rangle$ for all $\xi,\eta$. To see how $T$ behaves with respect to complex scalars,

$$\langle T(i\xi),\eta\rangle=\operatorname{Re}\langle\zeta,i\xi\otimes\eta\rangle=\operatorname{Re}\langle\zeta,\xi\otimes i\eta\rangle=\langle T\xi,i\eta\rangle=\langle -iT\xi,\eta\rangle.$$

As this can be done for all $\eta\in\mathcal{K}$, we get that $T(i\xi)=-iT\xi$. So $T(\lambda\xi)=\bar\lambda T\xi$; we say that $T$ is **conjugate-linear**. As $\zeta$ is real orthogonal to $H\otimes K$, if $\xi\in H$ and $\eta\in K$ then

$$\operatorname{Re}\langle T\xi,\eta\rangle=\operatorname{Re}\langle\zeta,\xi\otimes\eta\rangle=0.$$

This means that $TH\subset K_{\mathbb{R}}^{\perp}$ and $T^oK\subset H_{\mathbb{R}}^{\perp}$, where $T^o$ denotes the transpose of $T$ (or, to accommodate more to the terminology we have used for hundreds of pages, the "real adjoint"). Being a composition of two conjugate-linear maps, $T^oT:\mathcal{H}\to\mathcal{H}$ is actually linear and of course bounded. Now using that $\zeta$ is real orthogonal to $i(H_{\mathbb{R}}^{\perp}\otimes K_{\mathbb{R}}^{\perp})$, for $\xi\in H_{\mathbb{R}}^{\perp}$ and $\eta\in K_{\mathbb{R}}^{\perp}$

$$\operatorname{Re}\langle T\xi,i\eta\rangle=\langle\zeta,i(\xi\otimes\eta)\rangle=0.$$

Hence $TH_{\mathbb{R}}^{\perp}\subset(iK_{\mathbb{R}}^{\perp})_{\mathbb{R}}^{\perp}=iK_{\mathbb{R}}^{\perp\perp}=i\overline{K}$ and $T^o(iK_{\mathbb{R}}^{\perp})\subset(H_{\mathbb{R}}^{\perp})_{\mathbb{R}}^{\perp}=\overline{H}$. In particular

$$T^oTH_{\mathbb{R}}^{\perp}\subset T^o(i\overline{K})=-iT^o(\overline{K})\subset -iH_{\mathbb{R}}^{\perp}=iH_{\mathbb{R}}^{\perp}.$$

Iterating this, $(T^oT)^2H_{\mathbb{R}}^{\perp}\subset H_{\mathbb{R}}^{\perp}$. We can compute, for $\xi\in\mathcal{H}$,

$$\begin{aligned}\langle T^oT\xi,\xi\rangle&=\operatorname{Re}\langle T^oT\xi,\xi\rangle+i\operatorname{Im}\langle T^oT\xi,\xi\rangle\\&=\operatorname{Re}\langle T^oT\xi,\xi\rangle+i\operatorname{Re}\langle T^oT\xi,i\xi\rangle\\&=\|T\xi\|^2+i\operatorname{Re}\langle T\xi,-iT\xi\rangle\\&=\|T\xi\|^2+i\operatorname{Re}i\|T\xi\|^2=\|T\xi\|^2\geq 0.\end{aligned}$$

Thus $T^oT\in\mathcal{B}(\mathcal{H})^+$. This allows us to conclude, via functional calculus, that $T^oT$ is a limit of $p_n((T^oT)^2)$ for a sequence of real polynomials $p_n$ with $p_n(0)=0$. Therefore $T^oTH_{\mathbb{R}}^{\perp}\subset H_{\mathbb{R}}^{\perp}$. We then have that $T^oTH_{\mathbb{R}}^{\perp}\subset$

$H_{\mathbb{R}}^{\perp} \cap iH_{\mathbb{R}}^{\perp} = (H + iH)^{\perp} = \{0\}$. As $\ker T^o T = \ker T$, this means that $TH_{\mathbb{R}}^{\perp} = \{0\}$. Given $\eta \in \mathcal{K}$ and $\xi \in H_{\mathbb{R}}^{\perp}$,

$$\operatorname{Re}\langle T^o\eta, \xi\rangle = \operatorname{Re}\langle \eta, T\xi\rangle = \operatorname{Re}\langle \eta, 0\rangle = 0,$$

showing that $T^o\mathcal{K} \subset H_{\mathbb{R}}^{\perp\perp} = \overline{H}$. Then $T^o K \subset \overline{H} \cap H_{\mathbb{R}}^{\perp} = \{0\}$. Given $\xi \in \mathcal{H}$ and $\eta \in K$,

$$\operatorname{Re}\langle T\xi, \eta\rangle = \operatorname{Re}\langle \xi, T^o\eta\rangle = \operatorname{Re}\langle \xi, 0, \rangle = 0,$$

showing that $T\mathcal{H} \subset K_{\mathbb{R}}^{\perp}$. As $i\mathcal{H} = \mathcal{H}$, we also get that $T\mathcal{H} \subset iK_{\mathbb{R}}^{\perp}$. But then $T\mathcal{H} \subset K_{\mathbb{R}}^{\perp} \cap iK_{\mathbb{R}}^{\perp} = (K + iK)^{\perp} = \{0\}$. We have proven that $T = 0$, and this means that $\zeta = 0$. ☐

We recall that the environment for $\mathcal{M} \otimes \mathcal{N}$ is $\mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K})$.

***14.4.3. Theorem. (Commutant of the Tensor Product)***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ and $\mathcal{N} \subset \mathcal{B}(\mathcal{K})$ be von Neumann algebras. Then $(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}'$.*

*Proof.* Consider first the case where $\mathcal{M}$ and $\mathcal{N}$ both have cyclic vectors, say $\xi_0$ and $\eta_0$ respectively. Let $H = \mathcal{M}^{\mathrm{sa}}\xi_0$ and $K = \mathcal{N}^{\mathrm{sa}}\eta_0$. The assumption that $\xi_0$ and $\eta_0$ are cyclic means that $H + iH$ is dense in $\mathcal{H}$ and $K + iK$ is dense in $\mathcal{K}$. So $\overline{H \otimes K + i(H_{\mathbb{R}}^{\perp} \otimes K_{\mathbb{R}}^{\perp})} = \mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$ by Lemma 14.4.2. Lemma 14.4.1 gives us, with $\mathcal{A} = \mathcal{M}$, $\mathcal{B} = \mathcal{M}'$ and $\mathcal{A} = \mathcal{N}$, $\mathcal{B} = \mathcal{N}'$ respectively, that

$$H_{\mathbb{R}}^{\perp} = \overline{i\mathcal{M}'_{\mathrm{sa}}\xi_0}, \qquad K_{\mathbb{R}}^{\perp} = \overline{i\mathcal{N}'_{\mathrm{sa}}\eta_0}.$$

From this we get that $\mathcal{M}_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}_{\mathrm{sa}}\eta_0 + i(\mathcal{M}'_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}'_{\mathrm{sa}}\eta_0)$ is dense in $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$. Let $\zeta_0 = \xi_0 \otimes \eta_0$, which is cyclic for $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N}$. We have, since $\mathcal{M}_{\mathrm{sa}} \otimes \mathcal{N}_{\mathrm{sa}} \subset (\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})_{\mathrm{sa}}$ and similarly for the commutants,

$$\mathcal{M}_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}_{\mathrm{sa}}\eta_0 \subset (\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})_{\mathrm{sa}}\zeta_0,$$

$$\mathcal{M}'_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}'_{\mathrm{sa}}\eta_0 \subset (\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}')_{\mathrm{sa}}\zeta_0.$$

This implies that

$$\mathcal{M}_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}_{\mathrm{sa}}\eta_0 + i(\mathcal{M}'_{\mathrm{sa}}\xi_0 \otimes \mathcal{N}'_{\mathrm{sa}}\eta_0) \subset (\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})_{\mathrm{sa}}\zeta_0 + i(\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}')_{\mathrm{sa}}\zeta_0$$

and thus the latter is dense in $\mathcal{H} \mathbin{\overline{\otimes}} \mathcal{K}$. Then Lemma 14.4.1, with $\mathcal{A} = \mathcal{M} \otimes \mathcal{N}$ and $\mathcal{B} = \mathcal{M}' \otimes \mathcal{N}'$ gives us

$$(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})' = \mathcal{B}'' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}'.$$

When no cyclic vector is present, we note first that we always have $\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}' \subset (\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})'$ and $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N} \subset (\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}')'$. If we show that $(\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}')'$ commutes with $(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})'$ then we will have $(\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{N})' \subset (\mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}')'' = \mathcal{M}' \mathbin{\overline{\otimes}} \mathcal{N}'$ and we would be done. Because of Polarization, linearity, and continuity it is enough to show that

$$\langle XY(\xi\otimes\eta),\xi\otimes\eta\rangle=\langle YX(\xi\otimes\eta),\xi\otimes\eta\rangle \tag{14.13}$$

for all $X\in(\mathcal{M}'\,\overline{\otimes}\,\mathcal{N}')'$, $Y\in(\mathcal{M}\,\overline{\otimes}\,\mathcal{N})'$, and for all $\xi\in\mathcal{H}$ and $\eta\in\mathcal{K}$. Let $P\in\mathcal{M}'$ and $Q\in\mathcal{N}'$ the projections onto $\overline{\mathcal{M}\xi}$ and $\overline{\mathcal{N}\eta}$ respectively (Corollary 12.3.14). The von Neumann algebras $\mathcal{M}P\subset\mathcal{B}(P\mathcal{H})$ and $\mathcal{N}Q\subset\mathcal{B}(Q\mathcal{K})$ have cyclic vectors $\xi$ and $\eta$ respectively. So the first part of the proof applies and we have (using Proposition 14.1.1)

$$(\mathcal{M}P\,\overline{\otimes}\,\mathcal{N}Q)'=P\mathcal{M}'P\,\overline{\otimes}\,Q\mathcal{N}'Q \tag{14.14}$$

in $\mathcal{B}(P\mathcal{H}\,\overline{\otimes}\,Q\mathcal{K})$. Let $R=P\otimes Q\in\mathcal{M}'\,\overline{\otimes}\,\mathcal{N}'$. Then, using Proposition 14.1.1 and (14.14),

$$\begin{aligned}RX&=XR\in(\mathcal{M}'\,\overline{\otimes}\,\mathcal{N}')'R=(R(\mathcal{M}'\,\overline{\otimes}\,\mathcal{N}')R)'=(P\mathcal{M}'P\,\overline{\otimes}\,Q\mathcal{N}'Q)'\\&=(\mathcal{M}P\,\overline{\otimes}\,\mathcal{N}Q)''=\mathcal{M}P\,\overline{\otimes}\,\mathcal{N}Q.\end{aligned}$$

Also,

$$RYR\in R(\mathcal{M}\,\overline{\otimes}\,\mathcal{N})'R=(R(\mathcal{M}\,\overline{\otimes}\,\mathcal{N})R)'=(\mathcal{M}P\,\overline{\otimes}\,\mathcal{N}Q)'.$$

What this shows is that $RX$ and $RYR$ commute. Hence, as $RX=RXR$,

$$\begin{aligned}\langle XY(\xi\otimes\eta),\xi\otimes\eta\rangle&=\langle XYR(\xi\otimes\eta),R(\xi\otimes\eta)\rangle=\langle RXYR(\xi\otimes\eta),\xi\otimes\eta\rangle\\&=\langle RXRYR(\xi\otimes\eta),\xi\otimes\eta\rangle=\langle RYRXR(\xi\otimes\eta),\xi\otimes\eta\rangle\\&=\langle YX(\xi\otimes\eta),\xi\otimes\eta\rangle.\end{aligned}$$

Therefore (14.13) holds, and the proof is complete. □

With Theorem 14.4.3 at hand, the relations $(\mathcal{M}\otimes I_{\mathcal{K}})'=\mathcal{M}'\,\overline{\otimes}\,\mathcal{B}(\mathcal{K})$ and $(\mathcal{M}\,\overline{\otimes}\,\mathcal{B}(\mathcal{K}))'=\mathcal{M}\otimes I_{\mathcal{K}}$ from Exercise 13.4.27 are trivial to prove.

**14.4.1. Exercises.**

(14.4.1) Let $S,T\in\mathcal{B}(\mathcal{H})$ be selfadjoint. Show that $ST=TS$ if and only if $S\mathcal{H}\perp_{\mathbb{R}} iT\mathcal{H}$.

## 14.5. The Trace

A type I factor is $M_n(\mathbb{C})$ if finite-dimensional, or $\mathcal{B}(\mathcal{H})$ if infinite-dimensional. We have seen that in $M_n(\mathbb{C})$ there is a natural tracial state. On $\mathcal{B}(\mathcal{H})$ for infinite-dimensional $\mathcal{H}$ we considered as a special case the class of operators where the natural trace is finite (Section 10.7). But there is no tracial, linear, and bounded functional on $\mathcal{B}(\mathcal{H})$; this can be easily seen by using the unilateral shift (or, more abstractly Proposition 14.2.14) to get a projection $P$ such that $I_{\mathcal{H}} \sim P \sim I_{\mathcal{H}} - P$. If $\tau : \mathcal{B}(\mathcal{H}) \to \mathbb{C}$ is linear and tracial, then

$$\tau(P) = \tau(I_{\mathcal{H}}) = \tau(P) + \tau(I_{\mathcal{H}} - P) = 2\tau(P),$$

so $\tau(P) = 0$. For any finite projection $Q \in \mathcal{B}(\mathcal{H})$ we have that $I_{\mathcal{H}} - Q$ is infinite, so $0 = \tau(I_{\mathcal{H}} - Q) = -\tau(Q)$. Thus $\tau$ is zero on all projections. If $\tau$ is bounded, then $\tau = 0$ by Corollary 12.4.16. This should make it clear that tracial functionals belong in finite von Neumann algebras.

The property that distinguishes the trace from other functionals is, as mentioned in Definition 13.3.5, that $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. When $\mathcal{M}$ is not a factor, there are lots and lots of tracial functionals. For instance consider $\mathcal{M} = M_n(L^\infty[0,1])$. For any $\psi \in L^\infty[0,1]_* = L^1[0,1]$ the normal functional

$$\tilde{\psi}(T) = \sum_{k=1}^{n} \psi(T_{kk})$$

is tracial (Exercise 14.5.1). Early on researchers realized that the proper analog of the trace in non-factor algebras was a **centre-valued trace**. That is, a map that is linear, positive, and tracial, but that takes values in the centre of the algebra. Constructing this map is the main goal of this section. As mentioned, this can only work on finite von Neumann algebras, and in fact it characterizes them.

**14.5.1. Definition.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. A projection $P \in \mathcal{M}$ is* ***monic*** *if there exist $n \in \mathbb{N}$ and pairwise orthogonal projections $P_1, \dots, P_n \in \mathcal{M}$ with $P_k \sim P$ for all $k$ and $\sum_{k=1}^n P_k \in \mathcal{Z}(\mathcal{M})$.*

It is implies that the definition that a projection equivalent to a monic projection is itself monic.

A natural example of monic projection is any minimal projection in $M_n(\mathbb{C})$. For instance $E_{11}$ is monic because $\sum_k E_{kk} = I_n$ and $E_{kk} \sim E_{11}$ for all $k$. In $\mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$, the monic projections are precisely the infinite projections $P$ such that $I_{\mathcal{H}} - P$ is of the same rank (Exercise 14.5.2). We can do more with monic projections in non-atomic von Neumann algebras; concretely, we can always find a dyadic family of monic projections:

***14.5.2. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra with no abelian projections (that is, with no type I part). Then there exists a family*

$$\{P_{k,j} : \ k \in \mathbb{N}, \ j \in \{1, \dots, 2^k\}\} \subset \mathcal{M}$$

*of monic projections such that*

$$\sum_{j=1}^{2^k} P_{k,j} = I_{\mathcal{M}} \qquad k \in \mathbb{N},$$

$$P_{k,j} \sim P_{k,\ell}, \qquad k \in \mathbb{N},\ j, \ell \in \{1, \dots, 2^k\},$$

*and*

$$P_{k+1,2j-1} + P_{k+1,2j} = P_{k,j}, \qquad k \in \mathbb{N},\ j \in \{1, \dots, 2^k\}.$$

*Proof.* Let $P_{0,1} = I_{\mathcal{M}}$. By Proposition 14.2.22, there exist equivalent projections $P_{1,1}, P_{1,2} \in \mathcal{M}$ with $P_{1,1} + P_{1,2} = P_{0,1}$. Inductively, suppose we have $P_{k,1}, \dots, P_{k,2^k}$, pairwise equivalent, with $\sum_{j=1}^{2^k} P_{k,j} = I_{\mathcal{M}}$, and with $P_{k,2j-1} + P_{k,2j} = P_{k-1,j}$ for $j = 1, \dots, 2^{k-1}$. Let $V_1, \dots, V_{2^k}$ be partial isometries with

$$V_j^* V_j = P_{k,1}, \qquad V_j V_j^* = P_{k,j}.$$

We use Proposition 14.2.22 to get equivalent projections $P_{k+1,1}$ and $P_{k+1,2}$ in $\mathcal{M}$ with $P_{k+1,1} + P_{k+1,2} = P_{k,1}$. Then we define

$$P_{k+1,2j-1} = V_j P_{k+1,1} V_j^*, \qquad P_{k+1,2j} = V_j P_{k+1,2} V_j^*.$$

We have

$$\begin{aligned}
\sum_{j=1}^{2^{k+1}} P_{k+1,j} &= \sum_{j=1}^{2^k} P_{k+1,2j-1} + P_{k+1,2j} \\
&= \sum_{j=1}^{2^k} V_j (P_{k+1,1} + P_{k+1,2}) V_j^* \\
&= \sum_{j=1}^{2^k} V_j P_{k,1} V_j^* = \sum_{j=1}^{2^k} P_{k,j} = I_{\mathcal{M}}.
\end{aligned}$$

Also, from

$$P_{k+1,2j-1} = V_j P_{k+1,1} V_j^*$$

and

$$P_{k+1,1} V_j^* V_j P_{k+1,1} = P_{k+1,1} P_{k,1} P_{k+1,1} = P_{k+1,1}$$

we get $P_{k+1,2j-1} \sim P_{k+1,1}$ for all $j$. We similarly obtain $P_{k+1,2j} \sim P_{k+1,2} \sim P_{k+1,1}$ and therefore the projections $\{P_{k+1,j}\}$ are pairwise equivalent. $\square$

**14.5.3. Lemma.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a type $II_1$ von Neumann algebra, $\{P_{k,j}\}$ a dyadic family of monic projections as in Proposition 14.5.2, and $P \in \mathcal{M}$ a projection. Then there exist a nonzero projection $Z \in \mathcal{Z}(\mathcal{M})$ and $k, j$ such that $P_{k,j}Z \preceq PZ$.*

*Proof.* By working on $\mathcal{M}\,c(P)$ we may assume without loss of generality that $c(P) = I_{\mathcal{M}}$. A compression of a $\mathrm{II}_1$ von Neumann algebra by a central projection is again $\mathrm{II}_1$. If such $Z$ as in the statement did not exist, by Comparison (Theorem 14.2.8) we would have that $P \preceq P_{k,j}$ for all $k, j$. So for each $k$ there exists a projection $Q_k \in \mathcal{M}$ with $P \sim Q_k \leq P_{k,2^k-1}$. The projections $\{Q_k\}$ are pairwise orthogonal, for $P_{k+r,2^{k+r}-1} \leq P_{k,2^k}$ and so

$$P_{k,2^k-1}P_{k+r,2^{k+r}-1} = P_{k,2^k-1}P_{k,2^k}P_{k+r,2^{k+r}-1} = 0.$$

But now a usual index-shifting argument shows that, since $Q_k \sim P \sim Q_j$ for all $k, j$,

$$\sum_{k=1}^{\infty} Q_k \sim \sum_{k=2}^{\infty} Q_k,$$

contradicting that $\mathcal{M}$ is finite. □

Statement like that of Lemma 14.5.4 sometimes miss to mention that certain projections that add to a projection are pairwise orthogonal; this is always the case, though, as we showed in Proposition 10.5.5. We remark that in Lemma 14.5.4 below we do not require the projections to be pairwise equivalent.

**14.5.4. Lemma.**

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a finite von Neumann algebra. If $P \in \mathcal{M}$ is a projection, then there exist monic projections $\{P_j\} \subset \mathcal{M}$ with $P = \sum_j P_j$.*

*Proof.* We know from Theorem 14.3.2 that $\mathcal{M}$ is a direct sum of a type I algebra and a type $\mathrm{II}_1$ algebra; so we may consider the two cases separately. So either $\mathcal{M}$ is type $\mathrm{I}_n$ with $n \in \mathbb{N}$, or $\mathrm{II}_1$.

If we show that any projection admits a monic subprojection, then via Zorn's Lemma we can construct a maximal family $\{P_j\}$ of pairwise orthogonal monic subprojections of $P$. If $\sum_j P_j \neq P$, then we can use Lemma 14.5.3 to contradict the maximality. So we are left with the task of showing that any projection in $\mathcal{M}$ admits a monic subprojection. As we have no more need of the original $P$ in the Lemma, we will reuse the symbol $P$ for said arbitrary projection in $\mathcal{M}$.

When $\mathcal{M}$ is type $\mathrm{I}_n$, by Proposition 14.3.5 we know that $\mathcal{M} = M_n(\mathbb{C}) \otimes \mathcal{A}$ with $\mathcal{A}$ abelian. By Exercise 13.4.27, $\mathcal{Z}(\mathcal{M}) = I_n \otimes \mathcal{A}$. By Comparison (Theorem 14.2.8) there exists a projection $Z \in \mathcal{A}$ with

$$E_{11} \otimes Z = (E_{11} \otimes I_{\mathcal{A}})(I_n \otimes Z) \preceq P(I_n \otimes Z).$$

If such $Z$ is nonzero, then $E_{11} \otimes Z$ is monic (for $E_{11} \otimes Z \sim E_{kk} \otimes Z$ and $\sum_k E_{kk} \otimes Z = I_n \otimes Z \in \mathcal{Z}(\mathcal{M})$) and below $P$. Otherwise, we have $P \preceq E_{11} \otimes I_{\mathcal{A}}$. So there is $P \sim P' \leq E_{11} \otimes I_{\mathcal{A}}$. This gives, since $E_{11}$ is minimal in $M_n(\mathbb{C})$,

$$P' = (E_{11} \otimes I_{\mathcal{A}})P'(E_{11} \otimes I_{\mathcal{A}}) = \lambda E_{11} \otimes Y$$

for some $\lambda \in \mathbb{C}$ and $Y \in \mathcal{A}$. From $P'$ being a projection we get that $\lambda = 1$ and $Y$ is a projection. Then $P'$ is monic and equivalent to $P$.

When $\mathcal{M}$ is type $\mathrm{II}_1$, by Lemma 14.5.3 there exists a nonzero projection $Z \in \mathcal{Z}(\mathcal{M})$ and $k, j$ with $P_{k,j}Z \preceq PZ \leq P$. Then $P_{k,j}Z$ is a monic projection and a projection equivalent to it (hence, monic) is below $P$. □

We emphasize that the monic projections in Lemma 14.5.4 are not pairwise equivalent in general, so having infinitely many does not contradict that $\mathcal{M}$ is finite. For instance, with the dyadic projections we created in a $\mathrm{II}_1$ von Neumann algebra,

$$\sum_k P_{k,2^{k-1}+1} = I_{\mathcal{M}}.$$

**14.5.1. Existence of the Centre-Valued Trace.** We discussed conditional expectations starting on subsection 13.2.11.

***14.5.5. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra. Then there exists a normal conditional expectation $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$.*

*Proof.* As $\mathcal{Z}(\mathcal{M})$ is abelian, it is type I. Then so is $\mathcal{N} = \mathcal{Z}(\mathcal{M})'$ by Theorem 14.3.13. By Exercise 14.3.2 there exists an abelian projection $Q \in \mathcal{N}$ with $c(Q) = I_{\mathcal{N}}$. Since $\mathcal{M} \subset \mathcal{N}$, using Proposition 14.1.1 and that $\mathcal{Z}(\mathcal{N}) \subset \mathcal{Z}(\mathcal{M})$,

$$Q\mathcal{M}Q \subset Q\mathcal{N}Q = \mathcal{Z}(Q\mathcal{N}Q) = \mathcal{Z}(\mathcal{N})Q \subset \mathcal{Z}(\mathcal{M})Q. \tag{14.15}$$

Let $\theta : \mathcal{Z}(\mathcal{M}) \to \mathcal{Z}(\mathcal{M})Q$ be given by $\theta(T) = TQ$. This is a normal isomorphism by Corollary 14.1.4. Let $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ be given by $\mathcal{E}(T) = \theta^{-1}(QTQ)$; this is well-defined by (14.15). It is clear that $\mathcal{E}$ is linear and normal. It is also contractive, for

$$\|\mathcal{E}(T)\| = \|\theta^{-1}(QTQ)\| = \|QTQ\| \leq \|T\|.$$

And if $T \in \mathcal{Z}(\mathcal{M})$, then $QTQ = TQ$ and so $\mathcal{E}(T) = \theta^{-1}(\theta(T)) = T$. By Tomiyama's Theorem (Proposition 13.2.68), $\mathcal{E}$ is a conditional expectation onto $\mathcal{Z}(\mathcal{M})$. ☐

***14.5.6. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra and $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ a tracial conditional expectation. Then $\mathcal{M}$ is finite.*

*Proof.* Suppose that $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ is a tracial conditional expectation. Let $Z \in \mathcal{Z}(\mathcal{M})$ be the unique projection such that $Z\mathcal{M}$ is properly infinite and $(I_{\mathcal{M}} - Z)\mathcal{M}$ is finite (Proposition 14.2.13). By Proposition 14.2.14 there exist projections $P, Q \in \mathcal{M}$ with $P \sim Q \sim Z$ and $Z = P + Q$. Then

$$Z = \mathcal{E}(Z) = \mathcal{E}(P + Q) = \mathcal{E}(P) + \mathcal{E}(Q) = \mathcal{E}(Z) + \mathcal{E}(Z) = 2Z.$$

Hence $Z = 0$ and $\mathcal{M}$ is finite. ☐

The goal we are moving towards is to find, when $\mathcal{M}$ is finite, a conditional expectation as in Proposition 14.5.5 , but with the additional properties of faithfulness and traciality.

***14.5.7. Lemma.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ a conditional expectation. The following statements are equivalent:*

*(i)* $\mathcal{E}(ST) = \mathcal{E}(TS)$ *for all* $S, T \in \mathcal{M}$;

*(ii)* $\mathcal{E}(T^*T) = \mathcal{E}(TT^*)$ *for all* $T \in \mathcal{M}$;

*(iii)* $P \sim Q$ *implies* $\mathcal{E}(P) = \mathcal{E}(Q)$ *for projections* $P, Q \in \mathcal{M}$.

*Such $\mathcal{E}$ is necessarily faithful, and $\mathcal{M}$ is finite.*

*Proof.* (i) $\Longrightarrow$ (ii) Trivial.

(ii) $\Longrightarrow$ (iii) Trivial.

(iii) $\Longrightarrow$ (i) Given any projection $P$ and unitary $U$ in $\mathcal{M}$, we have $\mathcal{E}(UPU^*) = \mathcal{E}(P)$. Since $\mathcal{E}$ is bounded and any $T \in \mathcal{M}$ is a norm-limit of linear combinations of projections (Corollary 12.4.16), we get $\mathcal{E}(UTU^*) = \mathcal{E}(T)$. Applying this to $TU$ we get $\mathcal{E}(UT) = \mathcal{E}(TU)$. As any $S \in \mathcal{M}$ is a linear combination of unitaries, $\mathcal{E}(ST) = \mathcal{E}(TS)$.

Suppose that $T \in \mathcal{M}^+$ and $\mathcal{E}(T) = 0$. By Corollary 12.4.17 there exist $\lambda > 0$ and a spectral projection $P$ with $\lambda P \leq TP = PTP$. Then, since $\mathcal{E}$ is positive,

$$0 \le \mathcal{E}(P) \le \lambda^{-1}\mathcal{E}(TP) = \lambda^{-1}\mathcal{E}(T^{1/2}PT^{1/2}) \le \lambda^{-1}\mathcal{E}(T) = 0.$$

So $\mathcal{E}(P) = 0$. If $P \neq 0$, by Lemma 14.5.4 there exists a nonzero monic projection $Q \in \mathcal{M}$ with $Q \le P$, so $0 \le \mathcal{E}(Q) \le \mathcal{E}(P) = 0$. As $Q$ is monic, there exist projections $Q_1, \cdots, Q_n$ with $Q_k \sim Q$ for all $k$ and $\sum_k Q_k \in \mathcal{Z}(\mathcal{M})$. Then

$$\sum_k Q_k = \mathcal{E}\Big(\sum_k Q_k\Big) = \sum_k \mathcal{E}(Q_k) = \sum_k \mathcal{E}(Q) = 0.$$

This implies that $Q_k = 0$ for all $k$ and in particular $Q = 0$, a contradiction that implies that $T = 0$ and $\mathcal{E}$ is faithful.

Finally, from the existence of $\mathcal{E}$ we get that $\mathcal{M}$ is finite by Proposition 14.5.6. □

***14.5.8. Lemma.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a finite von Neumann algebra, $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ a normal conditional expectation, and $\varepsilon > 0$. Then there exists a projection $P \in \mathcal{M}$ with $\mathcal{E}(P) \neq 0$ and*

$$\mathcal{E}(TT^*) \le (1+\varepsilon)\mathcal{E}(T^*T), \qquad T \in P\mathcal{M}P. \tag{14.16}$$

*Proof.* We first construct a projection $Q_0 \in \mathcal{M}$ such that $\mathcal{E}$ is faithful on $Q_0\mathcal{M}Q_0$. For this we consider via Zorn's Lemma a maximal family $\{Q_j\}$ of pairwise orthogonal projections with $\mathcal{E}(Q_j) = 0$ for all $j$. Then we put $Q_0 = I_{\mathcal{M}} - \sum_j Q_j$. The normality of $\mathcal{E}$ guarantees $\mathcal{E}(Q_0) = I_{\mathcal{M}}$. If now $\mathcal{E}(Q_0T^*TQ_0) = 0$ for some $T$, if $Q_0T^*TQ_0 \neq 0$ then by Corollary 12.4.17 there exist $\lambda > 0$ and a spectral projection $P \in Q_0\mathcal{M}Q_0$ with $\lambda P \le (Q_0T^*TQ_0)P$. As $P$ commutes with $Q_0T^*TQ_0$ and $\mathcal{E}$ is positive, we get

$$\lambda\mathcal{E}(P) \le \mathcal{E}\big((Q_0T^*TQ_0)^{1/2}P(Q_0T^*TQ_0)^{1/2}\big) \le \mathcal{E}(Q_0T^*TQ_0) = 0.$$

So $\mathcal{E}(P) = 0$ and this would contradict the maximality of the family $\{Q_j\}$. That is, $\mathcal{E}$ is faithful on $Q_0\mathcal{M}Q_0$.

Next we use Zorn's Lemma to construct a maximal family $\{P_j, Q_j\} \subset Q_0\mathcal{M}Q_0$ where each of $\{P_j\}$ and $\{Q_j\}$ are pairwise orthogonal families of projections, with $P_j \sim Q_j$ and $\mathcal{E}(P_j) > \mathcal{E}(Q_j)$ for all $j$. If the family is empty this means that $P \sim Q$ implies $\mathcal{E}(P) = \mathcal{E}(Q)$ for all projections $P, Q \in \mathcal{M}$, and by Lemma 14.5.7 this would mean that $\mathcal{E}$ is tracial so (14.16) holds. If the family is nonempty and $\mathcal{E}$ is not tracial (for otherwise we already have (14.16)), we put $P' = Q_0 - \sum_j P_j$ and $Q' = Q_0 - \sum_j Q_j$. Then $P' \sim Q'$ by Proposition 14.2.7 and Exercise 14.2.25. From $\mathcal{E}(P_j) > \mathcal{E}(Q_j)$ for all $j$ we get $\mathcal{E}(P') < \mathcal{E}(Q')$. This implies that $Q' \neq 0$; and as $P' \sim Q'$, we also have $P' \neq 0$.

Let

$$c = \inf\{t > 0 : \ \mathcal{E}(R_1) \le t\,\mathcal{E}(R_2) \ \text{ for all } R_1, R_2 \in \mathcal{P}(Q_0\mathcal{M}Q_0),$$
$$R_1 \le P',\ R_2 \le Q',\ R_1 \sim R_2\}.$$

The number $c$ exists since the set is nonempty, for $P', Q'$ is one such pair. We have $c \le 1$ since $\mathcal{E}(R_1) \le \mathcal{E}(R_2)$ for any pair of equivalent projections $R_1 \le P'$ and $R_2 \le Q'$ (otherwise we would contradict the maximality of the family). It is also clear that $c$ is attained (note that each $t$ in the set works for all pairs of projections). If $c = 0$, we would get $\mathcal{E}(P') = 0$, which is impossible for $\mathcal{E}$ is faithful on $Q_0\mathcal{M}Q_0$. So $c > 0$. This means that there exist nonzero projections $R_1, R_2 \in Q_0\mathcal{M}Q_0$ with $R_1 \le P'$, $R_2 \le Q'$, $R_1 \sim R_2$ and

$$(1+\varepsilon)\mathcal{E}(R_1) \not\le c\,\mathcal{E}(R_2). \tag{14.17}$$

As $\mathcal{E}(R_1)$ and $\mathcal{E}(R_2)$ are positive and live in the abelian algebra $\mathcal{Z}(\mathcal{M})$, we can take

$$Z = 1_{(-\infty,0]}\big(c\,\mathcal{E}(R_2) - (1+\varepsilon)\mathcal{E}(R_1)\big) \in \mathcal{Z}(\mathcal{M}),$$

which will satisfy $Z \ne 0$ due to (14.17), and so

$$(1+\varepsilon)\mathcal{E}(R_1Z) > c\,\mathcal{E}(R_2Z). \tag{14.18}$$

Now consider a (possibly empty) maximal family $\{P_j, Q_j\} \subset \mathcal{M}$ with $\{P_j\}$ and $\{Q_j\}$ pairwise orthogonal families of projections, $P_j \le R_1Z$, $Q_j \le R_2Z$, and $P_j \sim Q_j$ for all $j$, and $(1+\varepsilon)\mathcal{E}(P_j) \le c\,\mathcal{E}(Q_j)$ for all $j$ (these are not the previous $P_j, Q_j$ in the proof). We form $P = R_1Z - \sum_j P_j$, $Q = R_2Z - \sum_j Q_j$ in $\mathcal{M}$. Since $R_1Z \sim R_2Z$, we get $P \sim Q$ from Proposition 14.2.7 and Exercise 14.2.25. By the normality of $\mathcal{E}$ and the inequality (14.18), we know that either $P \ne 0$ or $Q \ne 0$; but they are equivalent, so both $P, Q \ne 0$.

Given projections $P_1, P_2 \in P\mathcal{M}P$ with $P_1 \sim P_2$, because $P \sim Q$ we can find a projection $R \le Q$ with $P_1 \sim R$. Since $P_1, P_2 \le P \le R_1Z \le P'$ and $R \le Q \le R_2Z \le Q'$, we have

$$\mathcal{E}(P_1) \le c\,\mathcal{E}(R) \le (1+\varepsilon)\mathcal{E}(P_2)$$

(the first inequality by definition of $c$, and the second one because otherwise after cutting with a central projection we would be able to get the reverse inequality as in (14.18), and this would contradict the maximality of the family in the previous paragraph). Let $T, U \in P\mathcal{M}P$ with $U$ a unitary. Writing the operator $T^*T$ as a limit of positive-coefficient linear combinations of projections, and using that $U^*RU \sim R$ for any projection $R$ and that $\mathcal{E}$ is bounded, we get

$$\mathcal{E}(U^*T^*TU) \le (1+\varepsilon)\mathcal{E}(T^*T), \qquad T \in P\mathcal{M}P,\ U \in \mathcal{U}(P\mathcal{M}P).$$

Since $P\mathcal{M}P$ is finite, by Exercise 14.2.28 we can write $T^* = U|T^*|$ for a unitary $U \in P\mathcal{M}P$. Then

$$U^*T^*TU = |T^*|^2 = TT^*.$$

This way we get
$$\mathcal{E}(TT^*) \leq (1+\varepsilon)\mathcal{E}(T^*T), \qquad T \in P\mathcal{M}P,$$
as desired. □

**14.5.9. Proposition.**

*Let $\mathcal{M}$ be a finite von Neumann algebra and $\varepsilon > 0$. Then there exists a normal conditional expectation $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ such that*
$$\mathcal{E}(TT^*) \leq (1+\varepsilon)\mathcal{E}(T^*T), \qquad T \in \mathcal{M}. \tag{14.19}$$

*Proof.* Suppose for now that we can show that there exists a nonzero central projection $Z_0 \in \mathcal{Z}(\mathcal{M})$ such that there is a normal conditional expectation $\mathcal{E}_0 : \mathcal{M}Z_0 \to \mathcal{Z}(\mathcal{M})Z_0$ that satisfies (14.19) on $\mathcal{M}Z_0$. Then we can construct a maximal family $\{Z_j\}$ of pairwise orthogonal projections in $\mathcal{Z}(\mathcal{M})$ with (14.19) holding in $\mathcal{M}Z_j$ for all $j$. If $I_{\mathcal{M}} - \sum_j Z_j \neq 0$ we would contradict the maximality with our assumption. Then we construct
$$\mathcal{E}(T) = \sum_j \mathcal{E}_j(TZ_j).$$
The series converges sot by Exercise 12.1.10 and the fact that $\mathcal{E}_j(TZ_j) = Z_j\mathcal{E}_j(T)Z_j$. Given $\psi \in \mathcal{Z}(\mathcal{M})_*$ a normal state, we can define $\psi_j(T) = \psi(\mathcal{E}_j(TZ_j))$. This is normal and positive because $\mathcal{E}_j$ and $\psi$ are; that is, $\psi_j \in \mathcal{M}_*^+$. Then
$$\begin{aligned}\sum_j \|\psi_j\| = \sum_j \psi_j(I_{\mathcal{M}}) = \sum_j \psi(\mathcal{E}_j(Z_j)) = \sum_j \psi(Z_j)\\ = \psi\Big(\sum_j Z_j\Big) = \psi(I_{\mathcal{M}}) = 1.\end{aligned}$$
This implies that $\sum_j \psi_j$ converges in norm to a positive functional $\rho$, and $\rho$ is normal by Theorem 12.7.2 (norm limit of normal functionals is normal). Now
$$\psi(\mathcal{E}(T)) = \sum_j \psi(\mathcal{E}_j(TZ_j)) = \sum_j \psi_j(T) = \rho(T).$$
So $\psi \circ \mathcal{E}$ is normal. As this can be done for any $\psi$ normal (since any normal functional is a linear combination of normal states, Corollary 12.6.5), $\mathcal{E}$ is normal.

It remains to prove that the initial claim holds; that is, that there exists a nonzero central projection $Z_0 \in \mathcal{Z}(\mathcal{M})$ such that there is a normal conditional expectation $\mathcal{E}_0 : \mathcal{M}Z_0 \to \mathcal{Z}(\mathcal{M})Z_0$ that satisfies (14.19) on $\mathcal{M}Z_0$. From Proposition 14.5.5 and Lemma 14.5.8 we know that there exist a projection $P \in \mathcal{M}$ and a normal conditional expectation $\mathcal{E}_P$ onto $\mathcal{Z}(\mathcal{M})$ such that (14.16) holds. By Lemma 14.5.4 there exists a monic subprojection of $P$ in

$\mathcal{M}$, so we may assume without loss of generality that $P$ is monic. Hence there exist pairwise orthogonal projections $P_1, \ldots, P_n \in \mathcal{M}$ with $P_k \sim P$ for all $k$ and $Z_1 = \sum_k P_k \in \mathcal{Z}(\mathcal{M})$. Let $V_k \in \mathcal{M}$ be partial isometries with $V_k^* V_k = P_k$ and $V_k V_k^* = P$. This allows us to define $\mathcal{E}_0 : \mathcal{M}Z_1 \to \mathcal{Z}(\mathcal{M})Z_1$ as

$$\mathcal{E}_0(T) = \sum_{k=1}^{n} \mathcal{E}_P(V_k T V_k^*).$$

This satisfies, for $T \in \mathcal{M}Z_1$ and using $TT^* = TZ_1T^*$,

$$\begin{aligned}
0 \le \mathcal{E}_0(TT^*) &= \sum_{k=1}^{n} \mathcal{E}_P(V_k TT^* V_k^*) \\
&= \sum_{k,j=1}^{n} \mathcal{E}_P(V_k T P_j T^* V_k^*) = \sum_{k,j=1}^{n} \mathcal{E}_P(V_k T V_j^* V_j T^* V_k^*) \\
&\le \sum_{k,j=1}^{n} (1+\varepsilon)\mathcal{E}_P(V_j T^* V_k^* V_k T V_j^*) \\
&= (1+\varepsilon) \sum_{j=1}^{n} \mathcal{E}_P(V_j T^* T V_j^*) = (1+\varepsilon)\mathcal{E}_0(T^*T).
\end{aligned}$$

This almost makes the trick, but it $\mathcal{E}_0$ not a conditional expectation. It is linear and positive, it maps into $\mathcal{Z}(\mathcal{M})Z_10$, but $\mathcal{E}_0(Z_1) = nZ_1\mathcal{E}_P(P)$ is not $Z_1$. This we solve in the following way. As $\mathcal{E}_0(Z_1) \ne 0$ and positive (recall that $\mathcal{E}_P(P) \ne 0$ by Lemma 14.5.8), let

$$Z_0 = 1_{[\varepsilon,\infty)}(\mathcal{E}_0(Z_1)) \in \mathcal{Z}(\mathcal{M})Z_1, \qquad S = g(\mathcal{E}_0(Z_1)) \in \mathcal{Z}(\mathcal{M})Z_1,$$

where $g(t) = \frac{1}{t}\, 1_{[\varepsilon,\infty)}(t)$. Then $Z_0 = S\,\mathcal{E}_0(Z_1)$ and we define $\mathcal{E}(T) = S\,\mathcal{E}_0(T)$ for $T \in \mathcal{M}Z_0$. This still satisfies, for $T \in \mathcal{M}Z_0$,

$$\mathcal{E}(TT^*) = S\,\mathcal{E}_0(TT^*) \le (1+\varepsilon)\,S\,\mathcal{E}_0(T^*T) = (1+\varepsilon)\mathcal{E}(T^*T).$$

And for $T \in \mathcal{Z}(\mathcal{M})$,

$$\mathcal{E}(TZ_0) = S\mathcal{E}_0(TZ_0) = Z_0 T\, S\mathcal{E}_0(Z_0) = Z_0 T Z_0 = TZ_0.$$

Hence $\mathcal{E}$ is a linear and positive projection onto $\mathcal{Z}(\mathcal{M})Z$; by Exercise 14.5.3 it is a conditional expectation from $\mathcal{M}Z$ onto $\mathcal{Z}(\mathcal{M})Z$. □

We finally have all the ingredients needed.

**14.5.10. Theorem.**

*Let $\mathcal{M}$ be a von Neumann algebra. The following statements are equivalent:*

*(i) $\mathcal{M}$ is finite;*

*(ii) there exists a tracial conditional expectation $\mathcal{E} : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$.*

*When it exists, $\mathcal{E}$ is necessarily unique, faithful, and normal.*

*Proof.* (i) $\implies$ (ii) We assume that $\mathcal{M}$ is finite. By Proposition 14.5.9 we get normal conditional expectations $\mathcal{E}_n : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ with $\mathcal{E}_n(TT^*) \leq (1 + 1/n)\mathcal{E}_n(T^*T)$. It is tempting to extract a cluster point via Proposition 13.2.58 and it would certainly be a tracial conditional expectation, but it is not clear how to guarantee normality. But we will show that the sequence $\{\mathcal{E}_n\}$ is Cauchy in norm. For this, let us write $a_n = 1 + 1/n$ to simplify notation.

Let $P \in \mathcal{M}$ be a monic projection. So there exist projections $P_1, \ldots, P_r \in \mathcal{M}$ with $P_k \sim P_j$ for all $k, j$, and $Z = \sum_k P_k \in \mathcal{Z}(\mathcal{M})$. If $m \leq n$ then

$$\begin{aligned} \mathcal{E}_n(P) &\leq r^{-1}\, a_n \sum_{k=1}^r \mathcal{E}_n(P_k) = r^{-1}\, a_n \mathcal{E}_n(Z) = r^{-1}\, a_n Z \\ &= r^{-1}\, a_n \mathcal{E}_m(Z) = r^{-1}\, a_n \sum_{k=1}^r \mathcal{E}_m(P_k) \\ &\leq r^{-1}\, a_n\, r\, a_m \mathcal{E}_m(P) \leq a_m^2\, \mathcal{E}_m(P). \end{aligned}$$

Given a projection $P \in \mathcal{M}$, by Lemma 14.5.4 we can write $P = \sum_j P_j$ with $P_j$ monic for all $j$. Then, using the normality of the expectations,

$$\mathcal{E}_n(P) = \sum_j \mathcal{E}_n(P_j) \leq a_m^2 \sum_j \mathcal{E}_m(P_j) = a_m^2\, \mathcal{E}_m(P).$$

Given $T \in \mathcal{M}^+$, via the Spectral Theorem we can write $T$ as a norm-limit of linear combinations of projections, with all coefficients positive. It follows that $\mathcal{E}_n(T) \leq a_m^2 \mathcal{E}_m(T)$. That is, the operator $a_m^2 \mathcal{E}_m - \mathcal{E}_n$ is positive. From Propositions 13.2.22 and 13.2.24 we get

$$\|a_m^2 \mathcal{E}_m - \mathcal{E}_n\| = a_m^2 \mathcal{E}_m(I_{\mathcal{M}}) - \mathcal{E}_n(I_{\mathcal{M}}) = a_m^2 - 1.$$

Thus

$$\begin{aligned} \|\mathcal{E}_m - \mathcal{E}_n\| &\leq \|\mathcal{E}_m - a_m^2 \mathcal{E}_m\| + \|a_m^2 \mathcal{E}_m - \mathcal{E}_n\| \\ &\leq (a_m^2 - 1)\|\mathcal{E}_m\| + a_m^2 - 1 = 2(a_m^2 - 1). \end{aligned}$$

Therefore the sequence $\{\mathcal{E}_n\}$ is Cauchy in norm. With $\mathcal{E} = \lim_n \mathcal{E}_n$, it is a conditional expectation onto $\mathcal{Z}(\mathcal{M})$. Given $\psi \in \mathcal{M}_*$, we have that $\psi \circ \mathcal{E} = \lim_n \psi \circ \mathcal{E}_n$. As $\mathcal{M}_*$ is norm-closed, $\psi \circ \mathcal{E}$ is normal. This can be done for all $\psi \in \mathcal{M}_*$, so $\mathcal{E}$ is normal (see the argument in the proof of Corollary 12.6.12).

It remains to address the uniqueness. Suppose that $\Psi : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ is a tracial conditional expectation. Let $P \in \mathcal{M}$ be a monic projection. Then there exist pairwise equivalent projections $P_1, \dots, P_r \in \mathcal{M}$ with $Z = \sum_k P_k \in \mathcal{Z}(\mathcal{M})$. This gives

$$\begin{aligned}\Psi(P) &= r^{-1} \sum_{k=1}^{r} \Psi(P) = r^{-1} \sum_{k=1}^{r} \Psi(P_k) \\ &= r^{-1}\,\Psi(Z) = r^{-1}\,Z = r^{-1}\,\mathcal{E}(Z) = r^{-1} \sum_{k=1}^{r} \mathcal{E}(P_k) \\ &= \mathcal{E}(P).\end{aligned}$$

So $\Psi$ and $\mathcal{E}$ agree on monic projections. Given a projection $P \in \mathcal{M}$, by Lemma 14.5.4 there exist monic projections $\{P_j\}_{j\in J} \subset \mathcal{M}$ with $P = \sum_j P_j$. Using the normality of $\mathcal{E}$ and the positivity of $\Psi$,

$$\begin{aligned}\mathcal{E}(P) = \sum_j \mathcal{E}(P_j) &= \sum_j \Psi(P_j) = \lim_F \sum_{j\in F} \Psi(P_j) \\ &= \lim_F \Psi\Big(\sum_{j\in F} P_j\Big) \le \Psi(P).\end{aligned}$$

As we can do this for $I_{\mathcal{M}} - P$ and both $\mathcal{E}$ and $\Psi$ are unital, we get $\Psi(P) = \mathcal{E}(P)$. As both are bounded, an application of the Spectral Theorem shows that $\Psi(T) = \mathcal{E}(T)$ for all $T \in \mathcal{M}^{\mathrm{sa}}$, and hence $\Psi = \mathcal{E}$. In particular, $\Psi$ is faithful and normal.

(ii) $\Longrightarrow$ (i) This is Proposition 14.5.6 □

**14.5.11. Definition.**

*Let $\mathcal{M}$ be a finite von Neumann algebra. The unique tracial conditional expectation from Theorem 14.5.10 is called the* ***centre-valued trace****.*

When $\mathcal{M}$ is a factor, the centre-valued trace is a faithful normal tracial state. In the case of a type $\mathrm{I}_n$ factor it is the usual normalized trace. When $\mathcal{M}$ is a $\mathrm{II}_1$-factor this trace is usually denoted $\tau$.

***14.5.12. Corollary.***

*Let $\mathcal{M}$ be a finite von Neumann algebra and $\mathcal{T}$ its centre-valued trace. Let $P, Q \in \mathcal{M}$ be projections. Then $P \preceq Q$ if and only if $\mathcal{T}(P) \le \mathcal{T}(Q)$. In particular, $P \sim Q$ if and only if $\mathcal{T}(P) = \mathcal{T}(Q)$.*

*Proof.* Exercise 14.5.4. □

**14.5.13. Corollary.**

*Let $\mathcal{M}$ be a finite von Neumann algebra with centre-valued trace $\mathcal{T}$, and let $\varphi \in \mathcal{M}_*$ be a normal tracial state. Then there exists a state $\psi \in \mathcal{Z}(\mathcal{M})_*$ such that $\varphi = \psi \circ \mathcal{T}$.*

*Proof.* We define, for $T \in \mathcal{Z}(\mathcal{M})$, $\psi(T) = \varphi(T)$. Now the result follows from the fact that a normal tracial state is determined by its valued on the centre (Exercise 14.5.5). $\square$

**14.5.2. Dixmier's Property.** The proof of the existence of the centre-valued trace shown above follows mostly the ideas in [**KR97b**, Section 8.2]. There is also a very nice argument, due to Yeadon [**Yea71**], that constructs a "trace-like" map $\mathcal{Z}(\mathcal{M})_* \to \mathcal{M}_*$ (that is, showing that normal functionals on the centre have a unique extension satisfying a certain property) and defines the trace as the dual of this map; this proof is nicely presented in [**Dav25**]. The original proof, though, was due to Dixmier [**Dix49**] using entirely different ideas that we will mention here for they have separate value. Here is a long, incomplete, list of papers that study/use the Dixmier Property and similar averaging procedures: [**AGR23**, **APT73**, **APT73**, **Arc77**, **Arc79**, **AM08**, **ART17**, **Cho79**, **Con70**, **dlHS94**, **Gol56**, **Haa15**, **HN91**, **HKW86**, **KLR67**, **KR67**, **Ken20**, **Mag08**, **NRS16**, **NS16**, **Oza13**, **Pop99**, **Pop00**, **Pow75**, **Rie82**, **Rie83**, **Rin84**, **Sko16**, **SZ73**, **Zsi74**].

**14.5.14. Definition.**

*Let $\mathcal{M}$ be a von Neumann algebra and $T \in \mathcal{M}$. The* ***Dixmier orbit*** *of $T$ is the set*
$$\mathcal{D}_{\mathcal{M}}(T) = \overline{\operatorname{conv}}\{UTU^* : \ U \in \mathcal{U}(\mathcal{M})\}.$$
*We define $\mathcal{D}_{\mathcal{M}}$ to be the set of maps $\mathcal{M} \to \mathcal{M}$ of the form $T \longmapsto \sum_{k=1}^r t_k U_k T U_k^*$, where $t_1, \ldots, t_r$ are convex coefficients and $U_k \in \mathcal{U}(\mathcal{M})$ for all $k$.*

By definition
$$\mathcal{D}_{\mathcal{M}}(T) = \overline{\{\gamma(T) : \ \gamma \in \mathcal{D}_{\mathcal{M}}\}}.$$
The definition works as well for C$^*$-algebras, but we will only consider the von Neumann case here.

**14.5.15. Theorem.** *(Dixmier Property)*

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a von Neumann algebra and $T \in \mathcal{M}$. Then*
$$\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M}) \neq \varnothing.$$

*Proof.* Suppose first that $T = T^*$ and $\|T\| = 1$. Let $P = 1_{[0,\infty)}(T)$ and $Q = I_{\mathcal{M}} - P$. By Comparison there exists a central projection $Z \in \mathcal{Z}(\mathcal{M})$ such that $ZQ \preceq ZP$ and $(I_{\mathcal{M}} - Z)P \preceq (I_{\mathcal{M}} - Z)Q$. So there exists $P_1 \leq ZP$ with $P_1 \sim ZQ$; we put $P_2 = ZP - P_1$. Similarly there exists $Q_1 \sim (I_{\mathcal{M}} - Z)P$ with $Q_1 \leq (I_{\mathcal{M}} - Z)Q$ and we put $Q_2 = (I_{\mathcal{M}} - Z)Q - Q_1$. This means we have partial isometries $V, W \in \mathcal{M}$ with

$$V^*V = ZQ, \quad VV^* = P_1, \quad W^*W = (I_{\mathcal{M}} - Z)P, \quad WW^* = Q_1.$$

The six projections $ZQ$, $P_1$, $P_2$, $(I_{\mathcal{M}} - Z)P$, $Q_1, Q_2$ are pairwise orthogonal (note that $PQ = 0$). This guarantees that

$$U = V + V^* + W + W^* + P_2 + Q_2 \in \mathcal{M}$$

is a selfadjoint unitary that satisfies

$$U^*P_1U = ZQ, \qquad U^*ZQU = P_1, \qquad U^*P_2U = P_2,$$

$$U^*Q_1U = (I_{\mathcal{M}} - Z)P, \qquad U^*(I_{\mathcal{M}} - Z)U = Q_1, \qquad U^*Q_2U = Q_2.$$

By definition of $P$ and $Q$ we have $-ZQ \leq ZT \leq ZP = P_1 + P_2$. Conjugating this with $U$, $-P_1 \leq UTU^*Z \leq ZQ + P_2$. Averaging the two inequalities,

$$-\frac{1}{2}(ZQ + P_1) \leq \frac{1}{2}(ZT + UTU^*Z) \leq \frac{1}{2}(P_1 + ZQ) + P_2.$$

We also have $Z = Z(Q + P) = ZQ + P_1 + P_2$. Then $Z \geq ZQ + P_1$ and we get

$$-\frac{1}{2}Z \leq \frac{1}{2}(ZT + UTU^*Z) \leq Z.$$

We immediately get

$$-\frac{3}{4}Z \leq \frac{1}{2}(ZT + UTU^*Z) - \frac{1}{4}Z \leq \frac{3}{4}Z.$$

An analog computation that begins with $-(Q_1 + Q_2) = -(I_{\mathcal{M}} - Z)Q \leq (I_{\mathcal{M}} - Z)T \leq (I_{\mathcal{M}} - Z)P$ produces

$$-\frac{3}{4}(I_{\mathcal{M}} - Z) \leq \frac{1}{2}((I_{\mathcal{M}} - Z)T + UTU^*(I_{\mathcal{M}} - Z)) + \frac{1}{4}(I_{\mathcal{M}} - Z) \leq \frac{3}{4}(I_{\mathcal{M}} - Z).$$

Averaging the two sets of inequalities leads to

$$\left\|\frac{1}{2}(T + UTU^*) - \frac{1}{4}(2Z - I_{\mathcal{M}})\right\| \leq \frac{3}{4}. \tag{14.20}$$

The point of this is that we found a convex combination of unitary conjugates of $T$ and a selfadjoint central element such that their distance is less than 1.

We now consider a general $T$. Writing $T = T_1 + iT_2$, with $T_1, T_2$ selfadjoint, we can see (14.20) as saying that there is $\gamma_1 \in \mathcal{D}_{\mathcal{M}}$ and $Z_1 \in \mathcal{Z}(\mathcal{M})^{\rm sa}$ such that $\|\gamma_1(T_1) - Z_1\| \leq \frac{3}{4}\|T_1\|$. Next we apply (14.20) to $\gamma_1(T_1) - Z_1$ to get $\gamma_2' \in \mathcal{D}_{\mathcal{M}}$ and $Z_2' \in \mathcal{Z}(\mathcal{M})^{\rm sa}$ with $\|\gamma_2'(\gamma_1(T_1)) - \gamma_2'(Z_1) - Z_2'\| \leq \left(\frac{3}{4}\right)^2\|T_1\|$. Iterating and relabelling we get sequences $\{\gamma_n\} \subset \mathcal{D}_{\mathcal{M}}$ (where $\gamma_2 = \gamma_2' \circ \gamma_1$, etc.) and $\{Z_n\} \subset \mathcal{Z}(\mathcal{M})^{\rm sa}$ (where $Z_2 = \gamma_2'(Z_1) - Z_2'$, etc.) such that

$$\|\gamma_n(T_1) - Z_n\| \leq \left(\frac{3}{4}\right)^n\|T_1\|.$$

So if we fix $\varepsilon > 0$, we have shown that there exists $\gamma \in \mathcal{D}_{\mathcal{M}}$ and $Z \in \mathcal{Z}(\mathcal{M})^{\text{sa}}$ with $\|\gamma(T_1) - Z\| < \varepsilon$. If we apply this to the element $\gamma(T_2)$, we find $\beta \in \mathcal{D}_{\mathcal{M}}$ and $Y \in \mathcal{Z}(\mathcal{M})^{\text{sa}}$ with $\|\beta(\gamma(T_2)) - Y\| < \varepsilon$. Note that because $Z, Y$ are central, $\beta(Z) = Z$ and $\beta(Y) = Y$. Then

$$\begin{aligned} \|(\beta \circ \gamma)(T) - (Z + iY)\| &\leq \|\beta(\gamma(T_1)) - Z\| + \|\beta(\gamma(T_2)) - Y\| \\ &= \|\beta\big(\gamma(T_1) - Z\big)\| + \|\beta(\gamma(T_2)) - Y\| \\ &\leq \|\gamma(T_1) - Z\| + \|\beta(\gamma(T_2)) - Y\| \\ &< 2\varepsilon. \end{aligned}$$

So we can start with $T$ and produce a sequence $\{T_n\}$, where $T_{n+1} = \gamma_n(T_n)$ for some $\gamma_n \in \mathcal{D}_{\mathcal{M}}$, and a sequence $\{Z_n\} \subset \mathcal{Z}(\mathcal{M})$ with $\|T_n - Z_n\| \leq 2^{-n}$. From this we get

$$\begin{aligned} \|T_{n+1} - T_n\| \leq \|T_{n+1} - Z_n\| + \|Z_n - T_n\| &= \|\gamma_n(T_n - Z_n)\| + \|Z_n - T_n\| \\ \leq 2\|T_n - Z_n\| \leq 2^{-n+1}. \end{aligned}$$

It follows that

$$\|T_{n+k} - T_n\| \leq \sum_{j=n}^{n+k-1} \|T_{j+1} - T_j\| \leq \sum_{j=n}^{n+k-1} 2^{-j+1} \leq 2^{-n+2}.$$

That is, the sequence $\{T_n\}$ is Cauchy and is thus convergent. We also get

$$\|Z_{n+1} - Z_n\| \leq \|Z_{n+1} - T_{n+1}\| + \|T_{n+1} - T_n\| + \|T_n - Z_n\| \leq 2^{-n+2}$$

and therefore $\{Z_n\}$ is also Cauchy and converges to some $Z \in \mathcal{Z}(\mathcal{M})$. But then $T_n \to Z$, and henceforth $Z \in \mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})$. □

The Dixmier Property can be (and was) used to prove the existence of the centre-valued trace in a finite von Neumann algebra. But it doesn't save much effort, for one still needs some of the manipulations that led to Theorem 14.5.10. Regarding Corollary 14.5.16 below, we have found the implication (ii) $\implies$ (i) (properly, that the singleton Dixmier Property implies the existence of the centre-valued trace) mentioned in many papers and notes, but most seem to skim over the details of the proof. We hope that the details provided will be useful to the reader.

***14.5.16. Corollary.***

*Let $\mathcal{M}$ be a von Neumann algebra. The following statements are equivalent:*

*(i) $\mathcal{M}$ is finite;*

*(ii) $\mathcal{D}_{\mathcal{M}}(T) \cap \mathcal{Z}(\mathcal{M})$ consists of a single point for each $T \in \mathcal{M}$.*

*Proof.* (i) $\implies$ (ii) Fix $S \in \mathcal{M}$. Since $\mathcal{M}$ is finite, it admits a centre-valued trace $\mathcal{T}$ by Theorem 14.5.10. By Theorem 14.5.15 there exists $Z \in \mathcal{D}_\mathcal{M}(S) \cap \mathcal{Z}(\mathcal{M})$. Hence there exists $\{\gamma_n\} \subset \mathcal{D}_\mathcal{M}$ such that $\gamma_n(S) \to Z$. As $\mathcal{T}(\gamma(S)) = \mathcal{T}(S)$ for all $\gamma \in \mathcal{D}_\mathcal{M}$,

$$Z = \mathcal{T}(S) = \mathcal{T}(\lim_n \gamma_n(S)) = \lim_n \mathcal{T}(\gamma_n(S)) = \lim_n \mathcal{T}(S) = \mathcal{T}(S).$$

That is, $\mathcal{T}(S)$ is the unique element in $\mathcal{D}_\mathcal{M}(S) \cap \mathcal{Z}(\mathcal{M})$.

(ii) $\implies$ (i) For each $T \in \mathcal{M}$, let $\mathcal{E}(T) \in \mathcal{Z}(\mathcal{M})$ denote the unique element of $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M})$. This map is positive, for the all the operators in $\mathcal{D}_\mathcal{M}$ are positive. From Exercise 14.5.8 we have

$$\mathcal{D}_\mathcal{M}(S+T) \cap \mathcal{Z}(\mathcal{M}) \subset \overline{\mathcal{D}_\mathcal{M}(S) \cap \mathcal{Z}(\mathcal{M}) + \mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M})}.$$

By hypothesis these are singletons, so it follows that $\mathcal{E}(S+T) = \mathcal{E}(S) + \mathcal{E}(T)$ for all $S, T \in \mathcal{E}$. If $Z \in \mathcal{Z}(\mathcal{M})$ then $\mathcal{D}_\mathcal{M}(Z) = \{Z\}$ and hence $\mathcal{E}(Z) = Z$. By Exercise 14.5.3, $\mathcal{E}$ is a conditional expectation onto $\mathcal{Z}(\mathcal{M})$. Given $T \in \mathcal{M}$ and $U \in \mathcal{U}(\mathcal{M})$,

$$\mathcal{D}_\mathcal{M}(UT) = \mathcal{D}_\mathcal{M}(U^*(UT)U) = \mathcal{D}_\mathcal{M}(TU),$$

the first equality due to the fact that $\{UW : W \in \mathcal{U}(\mathcal{M})\} = \mathcal{U}(\mathcal{M})$. Then $\mathcal{E}(UT) = \mathcal{E}(TU)$. As every $S \in \mathcal{M}$ is a linear combination of unitaries, we get that $\mathcal{E}(ST) = \mathcal{E}(TS)$ for all $S, T \in \mathcal{M}$. Then $\mathcal{M}$ is finite by Proposition 14.5.6. ☐

*14.5.17. Remark.* When $\mathcal{M}$ is not finite the set $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M})$ might not be too interesting. For example, if $\mathcal{M} = \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$ for any $T \in \mathcal{K}(\mathcal{H})$ we have $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M}) = \{0\}$ (Exercise 14.5.10). Meanwhile, for $P \in \mathcal{B}(\mathcal{H})$ infinite projection with $I_\mathcal{H} - P$ infinite, $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M}) = [0,1]\, I_\mathcal{H}$ (Exercise 14.5.11).

**14.5.3. Weights.** When $\dim \mathcal{H} = \infty$ we know from Theorem 14.5.10 that $\mathcal{B}(\mathcal{H})$ does not have a trace. But what about the trace, as considered in Section 10.7? It was only defined for the trace-class operators, but not for all of $\mathcal{B}(\mathcal{H})$; not even for all $\mathcal{K}(\mathcal{H})$. On $\mathcal{B}(\mathcal{H})$ the trace is unbounded, defined on a $\sigma$-weak dense subspace, where it is linear, positive, and satisfies the tracial property. A similar situation can be found on a $\mathrm{II}_\infty$ factor $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$, where $\mathcal{M}$ is a $\mathrm{II}_1$-factor. In this case we have a unique faithful normal tracial state $\tau$ on $\mathcal{M}$, and we get a densely defined tracial map $\tau \otimes \mathrm{Tr}$ on $\mathcal{M} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{H})$.

**14.5.18. Definition.**

*Let $\mathcal{M}$ be a von Neumann algebra. A* ***weight*** *on $\mathcal{M}$ is a map $\psi : \mathcal{M}^+ \to [0,\infty]$ such that*

(i) $\psi(T+S) = \psi(T) + \psi(S), \quad T, S \in \mathcal{M}^+;$

*(ii)* $\psi(\lambda T) = \lambda\psi(T), \quad \lambda \geq 0, T \in \mathcal{M}^+$.

*The weight* $\psi$ *is*

*(i)* ***faithful*** *if* $\psi(T) = 0$ *implies* $T = 0$;

*(ii)* ***finite*** *if* $\psi(I_{\mathcal{M}}) < \infty$;

*(iii)* ***semifinite*** *if the* $*$*-subalgebra* $\operatorname{span}\{S \in \mathcal{M}^+ : \psi(S) < \infty\}$ *is* $\sigma$*-weakly dense in* $\mathcal{M}$;

*(iv)* ***normal*** *if for every non-decreasing net* $\{T_j\} \subset \mathcal{M}^+$, $\psi(\sup T_j) = \sup \psi(T_j)$;

*(v)* ***tracial*** *if* $\psi(T^*T) = \psi(TT^*)$ *for all* $T \in \mathcal{M}$;

A finite weight admits an obvious extension to a positive linear functional. The trace Tr on $\mathcal{B}(\mathcal{H})$ is an example of a faithful normal semifinite weight. One reason why weights are important is that not every von Neumann algebra has a faithful state (the existence of a faithful state forces the algebra to be $\sigma$-finite), but they all have faithful normal weights. A weight is always of the form $\sum_j \varphi_j$ with each $\varphi_j$ a state. The GNS construction can be performed for a weight, too. Here we will just consider the particular results we need.

***14.5.19. Proposition.***

*Let* $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ *be a factor. The following statements are equivalent:*

*(i)* $\mathcal{M}$ *is semifinite;*

*(ii)* $\mathcal{M}$ *admits a faithful normal semifinite tracial weight.*

*Proof.* (i) $\Longrightarrow$ (ii) If $\mathcal{M}$ is finite, then—being a factor—it has a faithful normal trace by Theorem 14.5.10. Otherwise, since $\mathcal{M}$ is semifinite, it is of type $\mathrm{II}_\infty$. By Proposition 14.3.9, we have $\mathcal{M} = \mathcal{N} \mathbin{\overline{\otimes}} \mathcal{B}(\mathcal{K})$ for a $\mathrm{II}_1$-factor $\mathcal{N}$ and a Hilbert space $\mathcal{K}$. As just mentioned, being finite, $\mathcal{N}$ admits a faithful normal tracial state $\tau$. Then $\tau \otimes \mathrm{Tr}$ is a faithful normal semifinite tracial weight on $\mathcal{M}$.

(ii) $\Longrightarrow$ (i) If $\mathcal{M}$ is finite, then it has a trace as just discussed. If $\mathcal{M}$ is type III then it is properly infinite and we can halve $I_{\mathcal{M}}$ to get projections $P, Q \in \mathcal{M}$ with $P \sim Q \sim I_{\mathcal{M}}$. But then

$$\tau(I_{\mathcal{M}}) = \tau(P + Q) = \tau(P) + \tau(Q) = 2\tau(I_{\mathcal{M}})$$

and then $\tau(I_{\mathcal{M}}) = 0$. As $\tau$ is positive, this means that $\tau = 0$, a contradiction. The only remaining possibility is that $\mathcal{M}$ is type $\mathrm{II}_\infty$, and hence semifinite. □

**14.5.20. Definition.**

*Let $\psi : \mathcal{M} \to [0, \infty]$ be a weight. We associate to $\psi$ the sets*

$$\mathcal{F}_\psi = \{T \in \mathcal{M} : \ \psi(T) < \infty\},$$
$$\mathcal{N}_\psi = \{T \in \mathcal{M} : \ \psi(T^*T) < \infty\},$$
$$\mathcal{M}_\psi = \mathcal{N}_\psi^* \mathcal{N}_\psi .$$

We have that $\mathcal{F}_\psi$ is a face in $\mathcal{M}$, $N_\psi$ is a left ideal in $\mathcal{M}$, $\mathcal{M}_\psi$ is a $*$-subalgebra, and $\mathcal{M}_\psi = \operatorname{span} \mathcal{F}_\psi$ (Exercise 14.5.12). The weight $\psi$ is semifinite precisely when $\mathcal{M}_\psi$ is $\sigma$-weakly dense in $\mathcal{M}$.

***14.5.21. Proposition.***

*Let $\mathcal{M}$ be a von Neumann algebra and $\psi$ a tracial weight on $\mathcal{M}$. The following statements are equivalent:*

*(i) $\psi$ is semifinite;*

*(ii) for each $P \in \mathcal{P}(\mathcal{M})$ there exists $Q \in \mathcal{P}(\mathcal{M})$ with $Q \leq P$ and $\psi(Q) < \infty$.*

*(iii) for each $T \in \mathcal{M}^+$ there exists $S \in \mathcal{M}^+$ with $S \leq T$ and $\psi(S) < \infty$.*

*Proof.* (i) $\implies$ (ii) Let $P \in \mathcal{P}(\mathcal{M})$. By hypothesis and Exercise 12.3.13 there exists a net $\{T_j\} \subset \mathcal{M}_\psi$ with $T_j \xrightarrow{\text{sot}} P$. Via Kaplansky's Density Theorem, since $P \geq 0$, we may assume that $T_j \geq 0$ and $\|T_j\| \leq 1$ for all $j$. As $\psi$ is tracial,

$$\psi(PT_jP) = \psi(T_j^{1/2} P T_j^{1/2}) \leq \psi(T_j) < \infty.$$

The convergence $PT_jP \xrightarrow{\text{sot}} P$ guarantees that for big enough $j$ we have $PT_jP \neq 0$. Fix such $j$ and write $T_0 = T_j$. Then $PT_0P \leq P$, nonzero, and $\psi(PT_0P) < \infty$. By Corollary 12.4.17 there exists $\lambda > 0$ and $Q \in \mathcal{P}(\mathcal{M})$ with $Q(PT_0P) = (PT_0P)Q$ and $\lambda Q \leq Q(PT_0P) \leq P$. Then $Q \leq P$ and $\psi(Q) < \infty$.

(ii) $\implies$ (iii) Let $T \in \mathcal{M}^+$. By Corollary 12.4.17 there exists $\lambda > 0$ and $P \in \mathcal{P}(\mathcal{M})$ with $TP = PT$ and $\lambda P \leq TP$. By hypothesis there exists a projection $Q \in \mathcal{M}^+$, nonzero, with $\psi(Q) < \infty$ and $Q \leq P$. Then

$$\lambda Q \leq \lambda P \leq TP = T^{1/2} P T^{1/2} \leq T.$$

(iii) $\implies$ (i) Given $P \in \mathcal{P}(\mathcal{M})$ by hypothesis there exists $S \in \mathcal{M}^+$ with $S \leq P$ and $\psi(S) < \infty$. Via the Spectral Theorem as in Corollary 12.4.17 there exists $\lambda > 0$ and a projection $Q \in \mathcal{M}$ that commutes with $S$ and with

$\lambda Q \leq QS$. Then

$$\lambda Q \leq QS = S^{1/2}QS^{1/2} \leq S \leq P.$$

This implies $Q \leq P$, for we get $(I_{\mathcal{M}} - P)Q = 0$. We have $\psi(Q) \leq \psi(S) < \infty$ Let $\{Q_j\}$ be a maximal family of pairwise orthogonal projections with $Q_j \leq P$ and $\psi(Q_j) < \infty$ for all $j$. We have $\sum_j Q_j = P$, for otherwise the above argument applied to $P - \sum_j Q_j$ would contradict the maximality. Then $P$ is the sot limit of finite sums of the $Q_j$, showing that $P \in \overline{\mathcal{M}_\psi}^{\text{sot}}$. As $\mathcal{M}$ is spanned in norm by its projections (Corollary 12.4.16), we have proved that $\overline{\mathcal{M}_\psi}^{\text{sot}} = \mathcal{M}$. Then $\overline{\mathcal{M}_\psi}^{\sigma\text{-weak}} = \mathcal{M}$ by Exercise 12.3.13. □

***14.5.22. Proposition.***

*Let $\mathcal{A} \subset \mathcal{M}$ be von Neumann algebras with the same unit, and with $\mathcal{A}$ abelian. Let $\tau$ be a normal tracial weight on $\mathcal{M}$ such that $\tau|_{\mathcal{A}}$ is semifinite. Then $\tau$ is semifinite.*

*Proof.* Let $\psi = \tau|_{\mathcal{A}}$. As $\psi$ is semifinite, by Proposition 14.5.21 there exists a nonzero projection $P \in \mathcal{A}$ with $\psi(P) < \infty$. Let $\{P_j\}$ be a maximal family of pairwise orthogonal projections in $\mathcal{A}$ such that $\psi(P_j) < \infty$ for all $j$. We necessarily have $\sum_j P_j = I_{\mathcal{A}} = I_{\mathcal{M}}$ for otherwise we could use Proposition 14.5.21 to contradict the maximality. Fix $T \in \mathcal{M}^+$. Then $T = S^*S$ for some $S \in \mathcal{M}$. By the normality of $\tau$,

$$\tau(T) = \tau(S^*S) = \tau\Big(\sum_j S^*P_jS\Big) = \sum_j \tau(S^*P_jS) = \sum_j \tau(P_jSS^*P_j).$$

If $\tau(T) \neq 0$ there exists $j$ with $\tau(P_jSS^*P_j) > 0$. We also have

$$\tau(P_jSS^*P_j) \leq \|S\|^2\tau(P_j) = \|S\|^2\, \psi(P_j) < \infty.$$

Then $S^*P_jS \leq T$ is nonzero with $\tau(S^*P_jS) < \infty$. When $\tau(T) = 0$, $T$ itself is below $T$ with $\psi(T)$ finite. By Proposition 14.5.21, $\tau$ is semifinite. □

The converse to Proposition 14.5.22 does not hold (Exercise 14.5.13).

**14.5.4. Exercises.**

(14.5.1) Let $\mathcal{M} = M_n(L^\infty[0,1])$ and $\psi \in L^\infty[0,1]_* = L^1[0,1]$. Show that the functional

$$\Psi(T) = \sum_{k=1}^n \psi(T_{kk})$$

is tracial.

(14.5.2) Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Show that a projection $P \in \mathcal{B}(\mathcal{H})$ is monic in the von Neumann algebra $\mathcal{B}(\mathcal{H})$ if and only if $\dim P\mathcal{H} = \dim(P\mathcal{H})^\perp$.

(14.5.3) Show that a map $\Psi : \mathcal{M} \to \mathcal{Z}(\mathcal{M})$ is a conditional expectation if and only if it is a linear and positive projection.

(14.5.4) Let $\mathcal{M}$ be a finite von Neumann algebra and $\mathcal{T}$ its centre-valued trace. Let $P, Q \in \mathcal{M}$ be projections. Show that $P \preceq Q$ if and only if $\mathcal{T}(P) \leq \mathcal{T}(Q)$.

(14.5.5) Let $\mathcal{M}$ be a finite von Neumann algebra and $\psi, \varphi \in \mathcal{M}_*$ two tracial normal states such that $\psi|_{\mathcal{Z}(\mathcal{M})} = \varphi|_{\mathcal{Z}(\mathcal{M})}$. Show that $\psi = \varphi$.

(14.5.6) Let $\mathcal{M}$ be a von Neumann algebra, $T_1, \ldots, T_r \in \mathcal{M}$, and $\varepsilon > 0$. Show that there exists $\gamma \in \mathcal{D}_\mathcal{M}$ and $Z_1, \ldots, Z_r \in \mathcal{Z}(\mathcal{M})$ such that

$$\|\gamma(T_k) - Z_k\| < \varepsilon, \qquad k = 1, \ldots, r.$$

(14.5.7) Let $\mathcal{M}$ be a von Neumann algebra and $T_1, \ldots, T_r \in \mathcal{M}$. Show that there exist $Z_1, \ldots, Z_r \in \mathcal{Z}(\mathcal{M})$ and $\{\gamma_n\} \subset \mathcal{D}_\mathcal{M}$ such that $\lim_n \gamma_n(T_k) = Z_k$, $k = 1, \ldots, r$.

(14.5.8) Let $\mathcal{M}$ be a von Neumann algebra and $S, T \in \mathcal{M}$. Show that

$$\mathcal{D}_\mathcal{M}(S+T) \cap \mathcal{Z}(\mathcal{M}) \subset \overline{\mathcal{D}_\mathcal{M}(S) \cap \mathcal{Z}(\mathcal{M}) + \mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M})}.$$

(14.5.9) Let $\mathcal{M}$ be a von Neumann algebra, $T \in \mathcal{M}$, $Z \in \mathcal{Z}(\mathcal{M})$. Show that

$$\mathcal{D}_\mathcal{M}(TZ) \cap \mathcal{Z}(\mathcal{M}) \subset \overline{Z\big(\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M})\big)}.$$

(14.5.10) Let $\mathcal{M} = \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$ and $T \in \mathcal{K}(\mathcal{H})^+$. Show that $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M}) = \{0\}$.

(14.5.11) Let $\mathcal{M} = \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} = \infty$ and $P \in \mathcal{B}(\mathcal{H})$ an infinite projection with $I_\mathcal{H} - P$ infinite. Show that $\mathcal{D}_\mathcal{M}(T) \cap \mathcal{Z}(\mathcal{M}) = [0,1]\, I_\mathcal{H}$.

(14.5.12) Given a weight $\psi : \mathcal{M}^+ \to [0, \infty]$, show that $\mathcal{F}_\psi$ is a face in $\mathcal{M}^+$, $N_\psi$ is a left ideal in $\mathcal{M}$, $\mathcal{M}_\psi$ is a $*$-subalgebra, and $\mathcal{M}_\psi = \operatorname{span} \mathcal{F}_\psi$.

(14.5.13) Show that the converse to Proposition 14.5.22 does not hold in general.

## 14.6. Examples of Factors

Examples of factors of all types $\mathrm{I}_n$, $\mathrm{I}_\infty$, $\mathrm{II}_1$, $\mathrm{II}_\infty$, and III were already known to Murray and von Neumann in the 1930s. In [vN40] examples of factors were constructed out of a crossed product construction. We will only consider the case where the group is discrete, as this is enough for our present needs.

**14.6.1. Crossed Products of von Neumann Algebras.**

**14.6.1. Definition.**

*Let $(\mathcal{M}, G, \alpha)$ be a $C^*$-dynamical system where $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ is a von Neumann algebra. The* ***von Neumann crossed product*** *is the von Neumann algebra $\mathcal{M} \overline{\rtimes}_\alpha G$ generated by $\mathcal{M} \rtimes_\alpha^r G$ in $\mathcal{B}(\mathcal{H} \mathbin{\overline{\otimes}} \ell^2(G))$.*

Because each $\alpha_g$ is an automorphism of $\mathcal{M}$, it is normal by Corollary 12.6.12. So the covariant relation

$$\tilde\pi(\alpha_g(T)) = U_g \tilde\pi(T) U_g^*,$$

holds for all $T \in \mathcal{M} \overline{\rtimes}_\alpha G$, where $U_g = I_{\mathcal{H}} \otimes \lambda_g$, $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{H})$ is the identity, and

$$\tilde\pi(T)(\xi \otimes \delta_g) = \alpha_g^{-1}(T)\xi \otimes \delta_g.$$

The conditional expectation from Proposition 13.5.10 was obtained as $\mathcal{E} = \mathrm{id}_{\mathcal{B}(\mathcal{H})} \mathbin{\overline{\otimes}} \tau$; and $\tau$ is normal because it is a point state, and faithful. We get that $\mathcal{E}$ is normal and faithful by Proposition 13.4.7.

The fact that our Hilbert space is $\mathcal{H} \otimes \ell^2(G)$ allows us to think of $\tilde T \in \mathcal{M} \overline{\rtimes}_\alpha G$ as block matrix indexed by $G$. Concretely, given $\xi, \eta \in \mathcal{H}$ and $g, h \in G$ we consider the form

$$[\xi, \eta] = \langle \tilde T(\xi \otimes \delta_h), \eta \otimes \delta_g \rangle, \qquad \xi, \eta \in \mathcal{H}.$$

It is straightforward to check that this is a sesquilinear form, with $|[\xi, \eta]| \leq \|\tilde T\|\,\|\xi\|\,\|\eta\|$. By Proposition 10.1.5 there exists $\tilde T_{g,h} \in \mathcal{B}(\mathcal{H})$ such that

$$[\xi, \eta] = \langle \tilde T_{g,h}\xi, \eta \rangle.$$

For $S \in \mathcal{M}'$ and $\tilde\pi(T)U_s \in (\tilde\pi \rtimes U)(\mathcal{M} \cdot G)$,

$$\begin{aligned}
\langle \big[\tilde\pi(T)U_r\big]_{g,h} S\xi, \eta \rangle &= \langle \tilde\pi(T)U_r(S\xi \otimes \delta_h), \eta \otimes \delta_g \rangle \\
&= \langle \tilde\pi(T)(S\xi \otimes \delta_{rh}), \eta \otimes \delta_g \rangle \\
&= \langle \tilde\pi(T)(S \otimes I_{\ell^2(G)})(\xi \otimes \delta_{rh}), \eta \otimes \delta_g \rangle \\
&= \langle (S \otimes I_{\ell^2(G)})\tilde\pi(T)(\xi \otimes \delta_{rh}), \eta \otimes \delta_g \rangle \\
&= \langle \tilde\pi(T)U_r(\xi \otimes \delta_h), S^*\eta \otimes \delta_g \rangle \\
&= \langle \big[\tilde\pi(T)U_r\big]_{g,h}\xi, S^*\eta \rangle = \langle S\big[\tilde\pi(T)U_r\big]_{g,h}\xi, \eta \rangle.
\end{aligned}$$

As this can be done for all $\xi, \eta \in \mathcal{H}$ and all $S \in \mathcal{M}'$, we get that $\big[\tilde\pi(T)U_r\big]_{g,h} \in \mathcal{M}'' = \mathcal{M}$. Via linearity and wot limits the equality holds for all $\tilde T \in \mathcal{M} \overline{\rtimes}_\alpha G$. That is, every element in the crossed product can be considered as a matrix indexed by $G$ and with entries in $\mathcal{M}$. In this notation we have $\mathcal{E}(\tilde T) = \tilde T_{e,e}$. This extends to

$$\tilde T_{g,h} = \mathcal{E}(U_g^* \tilde T U_h). \tag{14.21}$$

(Exercise 14.6.3).

We also have the "Toeplitz" relation

$$\tilde{T}_{g,h} = (U_g^* \tilde{T} U_g)_{e,g^{-1}h} \tag{14.22}$$

(Exercise 14.6.4). But we can get a direct relation between the entries of $\tilde{T}$. From Exercise 14.6.3, combined with the covariance of $\mathcal{E}$,

$$\begin{aligned} \tilde{T}_{g,h} &= \mathcal{E}(U_g^* \tilde{T} U_h) = \mathcal{E}(U_h^* \, U_{hg^{-1}} \tilde{T} U_h) \\ &= \alpha_h^{-1}(\mathcal{E}(U_{hg^{-1}} \tilde{T})) = \alpha_h^{-1}(\tilde{T}_{gh^{-1},e}). \end{aligned} \tag{14.23}$$

The reason we use the word "Toeplitz" to describe this stems from the fact that Toeplitz matrices are those with constant diagonals. Here the diagonals are not precisely constant, but they are entirely determined by one of its entries. In other words, as a matrix $\tilde{T} \in \mathcal{M} \overline{\rtimes}_\alpha G$ is entirely determined by any single one of its rows or columns. We have mentioned similar terminology in page 972 and Exercise 13.3.11.

***14.6.2. Lemma.***

*Let $\tilde{T} \in \mathcal{M} \overline{\rtimes}_\alpha G$. Then*

*(i) $\tilde{T} \in \tilde{\pi}(\mathcal{M})'$ if and only if $\tilde{T}_{g,e} R = \alpha_g^{-1}(R) \tilde{T}_{g,e}$ for all $R \in \mathcal{M}$ and $g \in G$;*

*(ii) $\tilde{T} \in \lambda(G)'$ if and only if $\tilde{T}_{ghg^{-1},e} = \alpha_g(\tilde{T}_{h,e})$ for all $g, h \in G$.*

*Proof.*

(i) To see that two elements in $\mathcal{M} \overline{\rtimes}_\alpha G$ are equal, we can check that their entries are equal; and due to (14.23) it is enough to check on a single column. Given $R \in \mathcal{M}$, the condition $[\tilde{T}\tilde{\pi}(R)]_{g,e} = [\tilde{\pi}(R)\tilde{T}]_{g,e}$ is equivalent to

$$\langle \tilde{T}\tilde{\pi}(R)(\xi \otimes \delta_e), \eta \otimes \delta_g \rangle = \langle \tilde{\pi}(R)\tilde{T}(\xi \otimes \delta_e), \eta \otimes \delta_g \rangle, \ g \in G, \ \xi, \eta \in \mathcal{H}.$$

This can be rewritten as

$$\langle \tilde{T}(R\xi \otimes \delta_e), \eta \otimes \delta_g \rangle = \langle \tilde{T}(\xi \otimes \delta_e), \alpha_g^{-1}(R^*)(\eta \otimes \delta_g) \rangle,$$

which in turn is

$$\langle \tilde{T}_{g,e} R\xi, \eta \rangle = \langle \alpha_g^{-1}(R) \tilde{T}_{g,e} \xi, \eta \rangle.$$

Therefore $\tilde{T}\tilde{\pi}(R) = \tilde{\pi}(R)\tilde{T}$ if and only if $\tilde{T}_{g,e} R = \alpha_g^{-1}(R) \tilde{T}_{g,e}$.

(ii) Now we have, using (14.21),

$$[TU_g]_{h,e} = \mathcal{E}(U_h^* \tilde{T} U_g) = \tilde{T}_{h,g}$$

and

$$[U_g \tilde{T}]_{h,e} = \mathcal{E}(U_h^* U_g \tilde{T}) = \mathcal{E}(U_{g^{-1}h}^* \tilde{T}) = \tilde{T}_{g^{-1}h,e}.$$

Therefore

$$\begin{aligned}\tilde{T}U_g = U_g\tilde{T} &\iff \tilde{T}_{h,g} = \tilde{T}_{g^{-1}h,e},\ h \in G\\ &\iff \alpha_g^{-1}(\tilde{T}_{hg^{-1},e}) = \tilde{T}_{g^{-1}h,e},\ h \in G\\ &\iff \alpha_g^{-1}(\tilde{T}_{ghg^{-1},e}) = \tilde{T}_{h,e},\ h \in G.\end{aligned}$$

□

For the examples we are looking at, we will only consider abelian von Neumann algebras.

**14.6.3. Definition.**

*Let $(\mathcal{A}, G, \alpha)$ be a $C^*$-dynamical group where $\mathcal{A}$ is a an abelian von Neumann algebra. We say that $G$ **acts freely** on $\mathcal{A}$ if for each $g \in G \setminus \{e\}$, if $T \in \mathcal{A}$ satisfies $TS = \alpha_g(S)T$ for all $S \in \mathcal{A}$, then $T = 0$.*

*The action is **ergodic** if the equality $\alpha_g(T) = T$ for $T \in \mathcal{A}$ and all $g \in G$, forces $T \in \mathbb{C}\, I_{\mathcal{A}}$.*

***14.6.4. Proposition.***

*The von Neumann algebra $\tilde{\pi}(\mathcal{A})$ is maximal abelian in $\mathcal{A} \,\overline{\rtimes}_\alpha G$ if and only if $G$ acts freely on $\mathcal{A}$.*

*Proof.* Suppose first that the action is free. If $\tilde{T} \in \tilde{\pi}(\mathcal{A})' \cap (\mathcal{A} \,\overline{\rtimes}_\alpha G)$, by Lemma 14.6.2 we have that $\tilde{T}_{g,e}A = \alpha_g^{-1}(A)T_{g,e}$ for all $A \in \mathcal{A}$. As the action is free this means that $\tilde{T}_{g,e} = 0$ for all $g \neq e$. Therefore $\tilde{T}_{g,h} = \alpha_h^{-1}(\tilde{T}_{gh^{-1},e}) = 0$ unless $g = h$. This gives

$$\begin{aligned}\langle \tilde{T}(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle = \langle \tilde{T}_{g,h}\xi, \eta\rangle &= \delta_{g,h}\,\langle \alpha_h^{-1}(T_{e,e})\xi, \eta\rangle\\ &= \langle \alpha_h^{-1}(T_{e,e})\xi \otimes \delta_h, \eta \otimes \delta_g\rangle\\ &= \langle \tilde{\pi}(T_{e,e})(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle.\end{aligned}$$

By linearity and continuity, $\tilde{T} = \tilde{\pi}(T_{e,e}) \in \tilde{\pi}(\mathcal{A})$. So, using that $\mathcal{A}$ is abelian,

$$\tilde{\pi}(\mathcal{A}) \subset \tilde{\pi}(\mathcal{A})' \cap (\mathcal{A} \,\overline{\rtimes}_\alpha G) \subset \tilde{\pi}(\mathcal{A})$$

and so $\tilde{\pi}(\mathcal{A})$ is maximal abelian.

Conversely, if the action is not free then there exists $r \in G$ with $r \neq e$ and nonzero $T \in \mathcal{A}$ with $TS = \alpha_r(S)T$ for all $S \in \mathcal{A}$. We define an operator $\tilde{T}$ via

$$\langle \tilde{T}(\xi \otimes \delta_h), \eta \otimes \delta_g\rangle = \delta_{h^{-1},rg}\langle \alpha_{rg}^{-1}(T)\xi, \eta\rangle.$$

It is routine to show that the right-hand-side is a bounded sesquilinear form and so $\tilde{T} \in \mathcal{A} \,\overline{\rtimes}_\alpha G$ exists by Proposition 10.1.5. It is not in $\tilde{\pi}(\mathcal{A})$, for $\tilde{T}_{e,e} = 0$.

We have $\tilde{T}_{g,e} = \delta_{g,r^{-1}}T$. When $g \neq r^{-1}$ it is zero; and when $g = r^{-1}$,

$$\tilde{T}_{g,e}S = TS = \alpha_r^{-1}(S)T = \alpha_r^{-1}(S)\tilde{T}_{g,e}.$$

Hence the equality works for all $g \in G$ and all $S \in \mathcal{A}$, and so $\tilde{T} \in \tilde{\pi}(\mathcal{A})' \cap (\mathcal{A} \overline{\rtimes}_\alpha G)$. This implies that $\mathcal{A}$ is not maximal abelian. $\square$

Let us finally get to the stage where the technicalities say something about von Neumann algebras.

**14.6.5. Proposition.**

*Let $(\mathcal{A}, G, \alpha)$ with $\alpha$ free. The following statements are equivalent:*

*(i) $\mathcal{A} \overline{\rtimes}_\alpha G$ is a factor;*

*(ii) the action is ergodic.*

*Proof.* (i) $\implies$ (ii) The hypothesis is that the centre of the crossed product is trivial. If $A \in \mathcal{A}$ with $\alpha_g(A) = A$ for all $g$, then $U_g\tilde{\pi}(A)U_g^* = \tilde{\pi}(A)$ for all $g \in G$, so $\tilde{\pi}(A)$ is a scalar multiple of the identity, which forces the same for $A$. So the action is ergodic.

(ii) $\implies$ (i) We know from Proposition 14.6.4 that $\tilde{\pi}(\mathcal{A})$ is maximal abelian in the crossed product. Then $\mathcal{Z}(\mathcal{A} \overline{\rtimes}_\alpha G) \subset \tilde{\pi}(\mathcal{A})$. Any $\tilde{T}$ in the centre satisfies $U_g\tilde{T}U_g^*$ for all $g \in G$. As is an element of $\mathcal{A}$, the ergodicity forces $\tilde{T} \in \mathbb{C}\, I_{\mathcal{A} \overline{\rtimes}_\alpha G}$. So $\mathcal{A} \overline{\rtimes}_\alpha G$ is a factor. $\square$

**14.6.6. Definition.**

*Let $G$ be a group with an action $\alpha$ over a von Neumann algebra $\mathcal{M}$. A weight $\psi$ on $\mathcal{M}$ is **$G$-invariant** if $\psi \circ \alpha_g = \psi$ for all $g \in G$.*

The point of the whole section was to get here, to the situation where the type of the crossed product is determined by a property of $\mathcal{A}$.

**14.6.7. Theorem.**

*Let $X$ be a separable metric space, and $\mu$ a diffuse $\sigma$-finite Borel measure on $X$. Let $\mathcal{A} = L^\infty(X)$ and $G$ an infinite countable discrete group with an action $\alpha$ on $\mathcal{A}$ that is free and ergodic. Then*

*(i) $\mathcal{A} \overline{\rtimes}_\alpha G$ is a type $II_1$-factor if and only if there exists a faithful normal finite $G$-invariant state on $\mathcal{A}$;*

*(ii)* $\mathcal{A} \rtimes_\alpha G$ *is a type* $II_\infty$*-factor if and only if there exists a faithful normal semifinite and infinite* $G$*-invariant weight on* $\mathcal{A}$*;*

*(iii)* $\mathcal{A} \rtimes_\alpha G$ *is a type III factor if and only if there is no nonzero normal semifinite* $G$*-invariant weight on* $\mathcal{A}$*.*

*Proof.* We know from Proposition 14.6.5 that $\mathcal{A} \rtimes_\alpha G$ is a factor. From Proposition 14.6.4 we know that $\mathcal{A}$ is maximal abelian in $\mathcal{A} \rtimes_\alpha G$. And as mentioned in the discussion above, together with Exercise 14.6.5, $\mathcal{E} : \mathcal{A} \rtimes_\alpha G \to \mathcal{A}$ is a faithful normal $G$-equivariant conditional expectation.

We first see that $\mathcal{A} \rtimes_\alpha G$ cannot be type I. If it were, as it is a factor it would be isomorphic to $\mathcal{B}(\mathcal{H})$, with a diffuse masa $\mathcal{A}$ and a normal conditional expectation $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{A}$. But we saw in Example 13.2.71 that such an expectation cannot be normal. Hence $\mathcal{A} \rtimes_\alpha G$ is either type II or III.

If $\mathcal{A} \rtimes_\alpha G$ is type II by Proposition 14.5.19 there exists a faithful normal semifinite tracial weight $\tau$. The restriction to $\mathcal{A}$ is still a faithful normal semifinite tracial weight. And we have, for $A \in \mathcal{A}$ and $g \in G$,

$$\tau(\alpha_g(A)) = \tau(U_g A U_g^*) = \tau(U_g^* U_g A) = \tau(A),$$

so $\tau|_\mathcal{A}$ is $G$-invariant. When $\mathcal{A} \rtimes_\alpha G$ is $\mathrm{II}_1$, $\tau$ can be chosen finite via Theorem 14.5.10 and so $\tau|_\mathcal{A}$ is finite. And when $\mathcal{A} \rtimes_\alpha G$ is $\mathrm{II}_\infty$ then $\tau(I_{\mathcal{A} \rtimes_\alpha G}) = \infty$ and still holds in $\mathcal{A}$, so $\tau|_\mathcal{A}$ is infinite.

Now suppose that $\mathcal{A}$ admits a faithful normal semifinite $G$-invariant tracial weight $\psi$. We define $\tau$ on $\mathcal{A} \rtimes_\alpha G$ by $\tau(T) = \psi(\mathcal{E}(T))$. This is faithful and normal because both $\psi$ and $\mathcal{E}$ are. And $\tau$ is semifinite by Proposition 14.5.22. So we have to see that $\tau$ is tracial. By definition of the crossed product, the subalgebra $\mathcal{M}_0 = \operatorname{span}\{AU_g : \ A \in \mathcal{A}, \ g \in G\}$ is sot-dense in $\mathcal{A} \rtimes_\alpha G$. Let $\tilde{T} = \sum_g A_g U_g$, where finitely many $A_g$ are nonzero. We have

$$\begin{aligned}
\mathcal{E}(\tilde{T}^*\tilde{T}) &= \mathcal{E}\Big(\Big[\sum_g A_g U_g\Big]^* \sum_g A_g U_g\Big) = \mathcal{E}\Big(\sum_g U_g^* A_g^* \sum_g A_g U_g\Big) \\
&= \mathcal{E}\Big(\sum_{g,h} U_h^* A_h^* A_g U_g\Big) = \mathcal{E}\Big(\sum_{g,h} U_h^* A_h^* A_g U_h U_{h^{-1}g}\Big) \\
&= \mathcal{E}\Big(\sum_{g,h} \alpha_h^{-1}(A_h^* A_g) U_{h^{-1}g}\Big) = \sum_{g,h} \alpha_h^{-1}(A_h^* A_g)\mathcal{E}(U_{h^{-1}g}) \\
&= \sum_g \alpha_g^{-1}(A_g^* A_g).
\end{aligned}$$

Similarly,

$$\mathcal{E}(\tilde{T}\tilde{T}^*) = \mathcal{E}\Big(\sum_g A_g U_g \Big[\sum_g A_g U_g\Big]^*\Big) = \mathcal{E}\Big(\sum_g A_g U_g \sum_g U_g^* A_g^*\Big)$$

$$= \mathcal{E}\Big(\sum_{g,h} A_g U_g U_h^* A_h^*\Big) = \mathcal{E}\Big(\sum_{g,h} A_g U_{gh^{-1}} A_h^* U_{gh^{-1}}^* U_{gh^{-1}}\Big)$$
$$= \mathcal{E}\Big(\sum_{g,h} A_g \alpha_{gh^{-1}}(A_h^*) U_{gh^{-1}}\Big) = \sum_{g,h} A_g \alpha_{gh^{-1}}(A_h^*)\mathcal{E}(U_{gh^{-1}})$$
$$= \sum_g A_g A_g^*.$$

Thus, since $\psi$ is normal,

$$\tau(\tilde T^*\tilde T) = \psi(\mathcal{E}(\tilde T^*\tilde T)) = \sum_g \psi(A_g^* A_g) = \sum_g \psi(A_g A_g^*)$$
$$= \psi(\mathcal{E}(\tilde T\tilde T^*)) = \tau(\tilde T\tilde T^*).$$

Comparing $\tau\big((\tilde S + \tilde T)^*(\tilde S + \tilde T)\big)$ and $\tau\big((\tilde S + \tilde T)(\tilde S + \tilde T)^*\big)$ we deduce that $\tau(\tilde S\tilde T) = \tau(\tilde T\tilde S)$ for all $\tilde S, \tilde T \in \mathcal{M}_0$. Then, using the density of $\mathcal{M}_0$ and that $\tau$ is normal, we obtain that $\tau$ is tracial on $\mathcal{A} \,\overline{\rtimes}_\alpha G$. Therefore $\mathcal{A} \,\overline{\rtimes}_\alpha G$ is semifinite by Proposition 14.5.19. If $\psi$ is finite then $\tau$ is finite and thus $\mathcal{A} \,\overline{\rtimes}_\alpha G$ is $\mathrm{II}_1$. And if $\psi$ is infinite then $\mathcal{A} \,\overline{\rtimes}_\alpha G$ is $\mathrm{II}_\infty$ (if it were $\mathrm{II}_1$ this would force $\psi$ to be finite).

Finally, the type III case works by discarding the other cases. That is, if $\mathcal{A} \,\overline{\rtimes}_\alpha G$ is type III then the two previous cases imply that $\mathcal{A}$ admits no nonzero normal semifinite $G$-invariant tracial weight. And conversely, if $\mathcal{A}$ admits no such weight, then $\mathcal{A} \,\overline{\rtimes}_\alpha G$ cannot be type II. ☐

**14.6.2. Examples of Factors.** The main goal here is to get a type III example as it is the missing one, but the group measure construction produces rather easily examples of the other types too.

*14.6.8. Example.* **Type I.** We saw earlier in the chapter that $M_n(\mathbb{C})$ is type $\mathrm{I}_n$ and when $\dim \mathcal{H} = \infty$ then $\mathcal{B}(\mathcal{H})$ is type $\mathrm{I}_\infty$. We can also take a discrete group $G$ with $|G| = n$ and consider its natural left action on $\ell^\infty(G)$. This is clearly free and ergodic. Then $\ell^\infty(G) \,\overline{\rtimes}_\alpha G$ is a factor with minimal projections $\{\delta_g : \ g \in G\}$, so type $\mathrm{I}_n$.

*14.6.9. Example.* **Type $\mathrm{II}_1$.** We discussed a few examples of $\mathrm{II}_1$-factors in Remark 14.3.8. All said examples were of the form $L(G)$ for an icc countable group $G$, but not all $\mathrm{II}_1$ factors are of said form. Here is an example constructed according to Theorem 14.6.7. Let $\mathcal{A} = L^\infty(\mathbb{T}) \subset \mathcal{B}(\ell^2(\mathbb{T}))$. Fix $\theta \in \mathbb{R} \setminus \mathbb{Q}$ and let $\mathbb{Z}$ act on $\mathcal{A}$ by $\alpha_n(f)(z) = f(e^{2i\pi\theta n}z)$ (translation by $n\theta$). This action is free and ergodic (Exercise 14.6.6). As the Lebesgue measure is translation invariant, the linear functional $\psi(f) = \int_{\mathbb{T}} f\,dm$ is $G$-invariant. It is also normal (by Monotone

Convergence), positive (being an integral against a positive measure); and it is faithful because a nonzero nonnegative continuous function is nonzero on a set of positive measure. Then Theorem 14.6.7 guarantees that $L^\infty(\mathbb{T}) \overline{\rtimes}_\alpha \mathbb{Z}$ is a $\mathrm{II}_1$-factor.

The group measure construction has been heavily used to produce examples of $\mathrm{II}_1$-factors. As has been mentioned Popa's Deformation/Rigidity (started in [**Pop06a**, **Pop06b**, **Pop07**]) has been highly successful for many years at producing better and better understood examples of $\mathrm{II}_1$-factors, and answering questions related to ridigity (like "for which groups $G$ does $L^\infty(X) \overline{\rtimes}_\alpha G \simeq L^\infty(Y) \overline{\rtimes}_\beta H$ imply $G \simeq H$?". We mention [**Pop06a**, **PV14**, **PV10**, **IPV13**, **OP10**, **IPP08**, **CI10**, **Dri20**] as a short sample of these works.

*14.6.10. Example.* **Type $\mathrm{II}_\infty$.** We saw in Proposition 14.3.9 and the paragraph preceding it that a $\mathrm{II}_\infty$ factor is necessarily of the form $\mathcal{N} \overline{\otimes} \mathcal{B}(\mathcal{H})$ with $\mathcal{N}$ a $\mathrm{II}_1$-factor and $\dim \mathcal{H} = \infty$. They can still be obtained via the group measure construction, though. Let $\mathcal{A} = L^\infty(\mathbb{R}) \subset \mathcal{B}(\ell^2(\mathbb{R}))$ and let $Q$ act by translation, that is $\alpha_q(f)(t) = f(t+q)$. This action is free and ergodic (Exercise 14.6.7). The linear functional induced by the Lebesgue measure is positive, faithful, normal (by Monotone Convergence), semifinite (for it is finite on all functions supported on compact sets, and these functions are $\sigma$-weakly dense in $L^\infty(\mathbb{R})$), and infinite (for $m(\mathbb{R}) = \infty$). By Theorem 14.6.7, $\mathcal{A} \overline{\rtimes}_\alpha G$ is a $\mathrm{II}_\infty$-factor.

*14.6.11. Remark.* The examples above are particular cases of the following situation. Let $\Gamma$ be a second countable locally compact topological group, with left Haar measure $\mu$ (this is a Borel measure). Fix a subgroup $G \subset \Gamma$, countable and dense. Consider the action of $G$ on $L^\infty(\Gamma)$ by left translation. Then the action is free and ergodic, so $L^\infty(\Gamma) \overline{\rtimes}_\alpha G$ is a factor. And it is of

- type $\mathrm{I}_n$, if and only if $\Gamma$ is discrete, $|\Gamma| = n$, and $G = \Gamma$;
- type $\mathrm{II}_1$, if and only if $\Gamma$ is compact and not discrete;
- type $\mathrm{II}_\infty$, if and only if $\Gamma$ is neither compact nor discrete.

*14.6.12. Example.* **Type III.** Let $\mathcal{A} = L^\infty(\mathbb{R})$, and $G = \mathbb{Q} \times \mathbb{Q}^*$ (where $Q^* = \mathbb{Q} \setminus \{0\}$), with the operation $(p_1, q_1)(p_2, q_2) = (q_1 p_2 + p_1, q_1 q_2)$, acting by translation in the first coordinate and multiplication on the second one; that is, given $(p, q) \in \mathbb{Q} \times \mathbb{Q}$, we define $\alpha_{p,q} f(t) = f(qt + p)$. This operation guarantees that $\alpha_{(p_1,q_1)} \alpha_{(p_2,q_2)} = \alpha_{(p_1,q_1)(p_2,q_2)}$. Let $\psi$ be a $G$-invariant normal semifinite weight on $\mathcal{A}$. If $\psi(f) = \infty$ for all positive $f \in C_c(\mathbb{R})$, the normality would imply that $\psi(f) = \infty$ for all non-negative $f \in \mathcal{A}$. As $\psi$ is semifinite, this means that there exists $f \in C_c(\mathbb{R})^+$ with $\psi(f) < \infty$. If $K = \{f \geq c\}$ for a small enough $c > 0$, then $K$ is compact and $1_K \leq c^{-1} f$, so $\psi(1_K) < \infty$. The fact that $f$ is continuous implies—via the Intermediate Value Theorem—that $K$ contains some interval $[r, s]$ with $\psi(1_{[r,s]}) < \infty$. But $\psi$ is invariant under translation and scaling, so $d = \psi(1_{[r,s]})$ satisfies $\psi(1_{[a,b]}) = d$ for all $a < b$ in $\mathbb{R}$. In particular (using that $1_{\{1\}} = 0$ on $\mathcal{A}$)

$$d = \psi(1_{[0,2]}) = \psi(1_{[0,1]}) + \psi(1_{[1,2]}) = 2d,$$

so $d = 0$. Using this and normality we quickly reach the conclusion that $\psi = 0$. Then Theorem 14.6.7 guarantees that $\mathcal{A} \overline{\rtimes}_\alpha (\mathbb{Q} \times \mathbb{Q}^*)$ is a type III factor.

**14.6.3. A brief technical clarification.** We finish with a technical comment that we learnt from [SS08, Page 21]. The computations in this section can be made a bit more slicker if instead of just making statements about the coordinates $\tilde{T}_{g,h}$ and/or approximating, one writes for example a product $\tilde{S}\tilde{T}$ as a series on the coefficients. Two of the main texts in the area do this ([KR97b, Proposition 13.1.2] and [Tak02, Proposition V.7.6]). But in general said series do not converge in the usual topologies. This was first recognized in [Mer85].

**14.6.4. Exercises.**

(14.6.1) Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$, $\pi : \mathcal{M} \to \mathcal{B}(\mathcal{H})$ the identity representation and $G$ a group that acts on $\mathcal{M}$ via $\alpha$. Show that if $S \in \mathcal{M}'$, then $S \otimes I_{\ell^2(G)} \in \tilde{\pi}(\mathcal{M})'$.

(14.6.2) Let $\tilde{T} \in \mathcal{M} \overline{\rtimes}_\alpha G$ and $g, h, r, s \in G$. Show that $[U_r \tilde{T} U_s]_{g,h} = \tilde{T}_{r^{-1}g,sh}$.

(14.6.3) Let $\tilde{T} \in \mathcal{M} \overline{\rtimes}_\alpha G$ and $g, h \in G$. Show that $\tilde{T}_{g,h} = \mathcal{E}(U_g^* \tilde{T} U_h)$.

(14.6.4) Prove (14.22).

(14.6.5) Use (14.23) to show that the conditional expectation $\mathcal{E} : \mathcal{M} \overline{\rtimes}_\alpha G \to \mathcal{M}$ is faithful.

(14.6.6) Let $\theta$ be an irrational number and $\alpha$ the translation action as in Example 14.6.9. Show that $\alpha$ is free and ergodic.

(14.6.7) Show that the action of $\mathbb{Q}$ on $L^\infty(\mathbb{R})$ by translation is free and ergodic.

## 14.7. $\mathrm{II}_1$-Factors

We have already mentioned several properties of $\mathrm{II}_1$-factors, and even some of the research directions on them. The combination of being a factor and being $\mathrm{II}_1$ provides lots of information. Here we will collect a few technical results that showcase the kind of manipulations these algebras allow, and that hint at how two entire research areas consist in big part in studying them. This section is definitely not a study guide for $\mathrm{II}_1$-factors.

**14.7.1. Proposition.**

*Let $\mathcal{M}$ be a $II_1$-factor and $\tau$ a tracial state. Then*

*(i) $\tau$ is faithful and normal;*

*(ii) $\tau$ is the unique tracial state on $\mathcal{M}$;*

*(iii) for every $t \in [0,1]$ there exists $P \in \mathcal{P}(\mathcal{M})$ with $\tau(P) = t$;*

*(iv) $P \sim Q$ if and only if $\tau(P) = \tau(Q)$.*

*Proof.* We know from Exercise 14.3.6 that $\tau(\mathcal{P}(\mathcal{M})) = [0,1]$. All remaining properties follow directly from Theorem 14.5.10 and Corollary 14.5.12. □

From Corollary 12.3.12 we know that a factor cannot have non-trivial wot-closed ideals. But there is more than that:

**14.7.2. Proposition.**

*A $II_1$-factor is simple as a $C^*$-algebra.*

*Proof.* Let $\mathcal{M}$ be a $\mathrm{II}_1$-factor and $\mathcal{J} \subset \mathcal{M}$ a nonzero norm-closed ideal. Fix $T \in \mathcal{J}^+$. By Corollary 12.4.17 there exists a nonzero projection $P \in \mathcal{M}$ and $\lambda > 0$ with $P \leq \lambda T$. As ideals as hereditary (as shown after Definition 11.5.19), $P \in \mathcal{J}$. By Proposition 14.2.10 there exists a family of pairwise orthogonal projections $\{P_1, \ldots, P_n\}$ with $P_j \preceq P$ for all $j$ and $\sum_j P_j = I_{\mathcal{M}}$. The family is necessarily finite because $I_{\mathcal{M}}$ is finite. For each $j$ there exists a projection $Q_j \in \mathcal{M}$ with $P_j \sim Q_j$ and $Q_j \leq P$. Because $\mathcal{J}$ is hereditary, $Q_j \in \mathcal{J}$. If $V_j \in \mathcal{M}$ is a partial isometry with $V_j^*V_j = P_j$ and $V_jV_j^* = Q_j$, by Exercise 10.4.12 $V_j = Q_jV_jP_j \in \mathcal{J}$. Then $P_j = V_j^*V_j \in \mathcal{J}$, which implies that $I_{\mathcal{M}} \in \mathcal{J}$ and hence $\mathcal{J} = \mathcal{M}$. □

Another proof of Proposition 14.7.2 can be achieved via the Dixmier Property (Exercise 14.7.2).

Being von Neumann algebras, the sot plays a crucial role in $\mathrm{II}_1$-factors. But the existence of the trace allows us to reformulate the topology in a very useful intrinsic way, in terms of the 2-norm $\|X\|_2 = \tau(X^*X)^{1/2}$.

***14.7.3. Proposition.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a II$_1$-factor. The sot topology agrees on closed bounded sets with the topology induced by the $\|\cdot\|_2$-norm.*

*Proof.* As both topologies are linear, we may assume without loss of generality that we are working on the closed unit ball of $\mathcal{M}$ and that nets converge to 0. Let $\{X_j\} \subset \mathcal{M}$ with $\|X_j\| \leq 1$ for all $j$. Suppose first that $X_j \xrightarrow{\text{sot}} 0$. From $\|\,|X_j|\xi\| = \langle X_j^* X_j \xi, \xi\rangle = \|X_j\xi\|$ we get that $|X_j| \xrightarrow{\text{sot}} 0$. As $\tau$ is normal, $\tau(|X_j|) \to 0$. And then

$$\|X_j\|_2^2 = \tau(|X_j|^2) = \tau(|X_j|^{1/2}|X_j|\,|X_j|^{1/2}) \leq \tau(|X_j|) \to 0.$$

Conversely, suppose that $\|X_j\|_2 \to 0$. The GNS representation $\pi_\tau : \mathcal{M} \to \mathcal{B}(\mathcal{H}_\tau)$ is faithful by Proposition 14.7.2. By Exercise 12.6.5, $\pi_\tau$ is normal. For any $T \in \mathcal{M}$,

$$\begin{aligned}\|\pi_\tau(X_j)T\Omega_\tau\|^2 &= \tau(T^* X_j^* X_j T) = \tau(X_j T T^* X_j^*) \\ &\leq \|T\|^2\,\tau(X_j X_j^*) = \|T\|^2\,\tau(X_j^* X_j) = \|T\|^2\,\|X_j\|_2^2 \to 0.\end{aligned}$$

For arbitrary $\xi \in \mathcal{H}_\tau$ and $\varepsilon > 0$ there exists $T \in \mathcal{M}$ such that $\|\xi - T\Omega_\tau\| < \varepsilon$. Then

$$\|\pi_\tau(X_j)\xi\| \leq \|\pi_\tau(X_j)(\xi - T\Omega_\tau)\| + \|\pi_\tau(X_j)T\Omega_\tau\| \leq \varepsilon + \|\pi_\tau(X_j)T\Omega_\tau\|.$$

Hence $\limsup_j \|\pi_\tau(X_j)\xi\| \leq \varepsilon$. As $\varepsilon$ was arbitrary, the Limsup Routine guarantees that the limit exists and is zero. Hence $\pi_\tau(X_j) \xrightarrow{\text{sot}} 0$. This implies that $\pi_\tau(X_j) \xrightarrow{\sigma\text{-weak}} 0$ via Lemma 12.1.21. Since $\pi_\tau$ is an isomorphism and the $\sigma$-weak topology is intrinsic to $\mathcal{M}$ (being its weak$^*$-topology as a dual), we get that $X_j \xrightarrow{\sigma\text{-weak}} 0$ and therefore $X_j \xrightarrow{\text{sot}} 0$ by Lemma 12.1.21. ☐

In general the adjoint operation is not sot-continuous (Exercise 12.1.1). But, in a II$_1$-factor,

***14.7.4. Corollary.***

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a II$_1$ factor. The map $X \longmapsto X^*$ is sot-continuous on bounded sets of $\mathcal{M}$.*

*Proof.* We use Proposition 14.7.3. If $\{X_j\} \subset \mathcal{M}$ with $\|X_j\| \leq c$ for all $j$ and $X_j \xrightarrow{\text{sot}} 0$, then $\|X_j\|_2 \to 0$. Thus

$$\|X_j^*\|_2 = \tau(X_j X_j^*)^{1/2} = \tau(X_j^* X_j)^{1/2} = \|X_j\|_2 \to 0.$$

Using again Proposition 14.7.3, $X_j^* \xrightarrow{\text{sot}} 0$. ☐

### *14.7.5. Proposition.*

*Let $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ be a $II_1$-factor and $\{P_n\}, \{Q_n\} \subset \mathcal{P}(\mathcal{M})$ two decreasing sequences of projections such that $P_n \sim Q_n$ for all $n$. Then there exists a unitary $U \in \mathcal{M}$ such that $UP_nU^* = Q_n$ for all $n$.*

*Proof.* Let $\tau$ be the unique trace on $\mathcal{M}$. Using Proposition 14.7.1 twice we get first that $\tau(P_n) = \tau(Q_n)$ and then that $P_k - P_{k+1} \sim Q_k - Q_{k+1}$ since they have equal trace. For each $k \in \mathbb{N}$, let $W_k \in \mathcal{M}$ be a partial isometry with $W_k^* W_k = P_k - P_{k+1}$ and $W_k W_k^* = Q_k - Q_{k+1}$. In particular,

$$W_k = (Q_k - Q_{k+1}) W_k (P_k - P_{k+1}).$$

Let $W = \sum_k W_k$; this series converges sot by Exercise 12.1.10. Also because of the pairwise orthogonal ranges, $\|W\| \leq 1$ and the same is true for all the partial sums of the series. This allows us to apply Corollary 14.7.4 and Proposition 12.1.13 to get

$$W^*W = \sum_k W_k^* W_k = \sum_k P_k - P_{k+1} = P_1,$$

$$WW^* = \sum_k W_k W_k^* = \sum_k Q_k - Q_{k+1} = Q_1.$$

Also, for any $n \in \mathbb{N}$,

$$WP_nW^* = \sum_{k \geq n} W_k P_k W_k^* = \sum_{k \geq n} W_k W_k^* = \sum_{k \geq n} Q_k - Q_{k+1} = Q_n.$$

Finally, we have $I_{\mathcal{M}} - P_1 \sim I_{\mathcal{M}} - Q_1$ by Exercise 14.2.24. Let $V$ be a partial isometry with $V^*V = I_{\mathcal{M}} - P_1$, $VV^* = I_{\mathcal{M}} - Q_1$. Then $U = W + V$ is the desired unitary. □

The next result is folklore. We are not aware of where/if it might have appeared in print.

### *14.7.6. Corollary.*

*Let $M$ be a $II_1$-factor, and $\pi : M \to \mathcal{B}(\mathcal{H})$ an irreducible representation. Then $\mathcal{H}$ is non-separable.*

*Proof.* Since $M$ is simple, $\pi$ is faithful by Proposition 14.7.2. By hypothesis, $\pi(\mathcal{M})'' = \overline{\pi(\mathcal{M})}^{\text{sot}} = \mathcal{B}(\mathcal{H})$. This shows that $\pi(\mathcal{M})$ is not a von Neumann algebra, as it is not equal to its double commutant ($\pi(\mathcal{M})$ does not have minimal projections, while $\mathcal{B}(\mathcal{H})$ does). By Proposition 12.7.3 there exists a monotone net $\{Q_j\} \subset \mathcal{M}$ of projections such that its sot-limit $\tilde{Q} = \lim_{\text{sot}} \pi(Q_j) \notin \pi(\mathcal{M})$. As $\mathcal{M}$ is a von Neumann algebra, $Q = \lim_{\text{sot}} Q_j$ exists (Proposition 12.1.10). Both $\tilde{Q}$ and $Q$ are projections by Proposition 12.1.13. As $Q_j \leq Q$ for all $j$ we have $\pi(Q_j) \leq \pi(Q)$ for all $j$, and so $\tilde{Q} \leq \pi(Q)$; and these two cannot be equal for $\tilde{Q} \notin \pi(\mathcal{M})$.

Let $P_j = Q - Q_j$; then $P_j \xrightarrow{\text{sot}} 0$ in $\mathcal{M}$, while $\pi(P_j) \xrightarrow{\text{sot}} \pi(Q) - \tilde{Q} \neq 0$. As the trace $\tau$ is normal, $\tau(P_j) \to 0$. This allows us to choose, for each $n \in \mathbb{N}$, indices $j_n$ such that $P_{j_{n+1}} \leq P_{j_n}$ and $\tau(P_{j_{n+1}}) < \tau(P_{j_n})/2$. After relabelling these projections as $\{P_n\}_{n=0}^{\infty} \subset \mathcal{M}$ and $\tilde{P} = \pi(Q) - \tilde{Q}$, we have that $\tau(P_{n+1}) < \tau(P_n)/2$ for all $n$, the sequence $\{P_n\}$ is decreasing, and $\pi(P_n) \geq \tilde{P}$ for all $n$ with $\pi(P_n) \xrightarrow{\text{sot}} \tilde{P}$.

Let $P_{1,1} = P_1$. From $P_1 \leq P_0$ and $\tau(P_1) < \tau(P_0)/2$, using (iii) from Proposition 14.7.1 on the $\mathrm{II}_1$-factor $P_0\mathcal{M}P_0$ there exists $P_{1,2} \leq P_0$ with $P_{1,2}P_{1,1} = 0$ and a partial isometry $V_1$ with $V_1^*V_1 = P_{1,1}$ and $V_1V_1^* = P_{1,2}$. Inductively, for each $n$ we set $P_{n,1} = P_n$ and we construct $P_{n,2}$ with $P_{n,1}P_{n,2} = 0$ and a partial isometry $V_n$ such that $V_n^*V_n = P_{n,1}$ and $V_nV_n^* = P_{n,2}$. Now we build dyadic projections as follows:

$$\begin{aligned}
P_{n,3} &= V_{n-1}P_{n,1}V_{n-1}^*, & & P_{n,4} = V_{n-1}P_{n,2}V_{n-1}^* \\
P_{n,5} &= V_{n-2}P_{n,1}V_{n-2}^*, & \ldots, \quad & P_{n,8} = V_{n-2}P_{n,4}V_{n-2}^* \\
P_{n,9} &= V_{n-3}P_{n,1}V_{n-3}^*, & \ldots, \quad & P_{n,16} = V_{n-3}P_{n,8}V_{n-3}^* \\
& \vdots \\
P_{n,2^{n-1}+1} &= V_1P_{n,1}V_1^*, & \ldots, \quad & P_{n,2^n} = V_1P_{n,2^n}V_1^*.
\end{aligned}$$

So we get projections $\{P_{n,k}\} \subset \mathcal{M}$, with $n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^n\}$, where $P_{n+1,2k-1} + P_{n,2k} \leq P_{n,k}$, $P_{n,1} = P_n$, and $P_{n,k} \sim P_n$ for all $k \in \{1, \ldots, 2^n\}$.

Given any $t \in [0,1]$ we can write it as a decreasing limit of dyadic numbers, $t = \lim_k m_k/2^k$, where $\{m_k\}$ is a sequence of integers with $m_k \geq 2m_{k+1}$. This way we can uniquely associate to $t$ the decreasing sequence of projections $\{P_{k,m_k}\}$. So there exists a projection $\tilde{P}_t$ such that $\pi(P_{k,m_k}) \xrightarrow{\text{sot}} \tilde{P}_t$. Using Proposition 12.1.13 we get that for $s \neq t$, $\tilde{P}_s\tilde{P}_t = 0$, since for large enough $k$ we have $m_k^s \neq m_k^t$ and hence $\pi(P_{k,m_k^s})\pi(P_{k,m_k^t}) = 0$ for all $k$ big enough. If we show that $\tilde{P}_t \neq 0$ for all $t$, we would have constructed an uncountable family of pairwise orthogonal projections, showing that $\mathcal{H}$ is non-separable.

So it remains to see that $\tilde{P}_t \neq 0$ for all $t \in [0,1]$. We know that $\tilde{P}_0 \neq 0$. Fix a sequence $\{P_{k,m_k}\} \subset \mathcal{M}$ with $\pi(P_{k,m_k}) \xrightarrow{\text{sot}} \tilde{P}_t$. By Proposition 14.7.5 there exists a unitary $U \in \mathcal{M}$ with $UP_{k,1}U^* = P_{k,m_k}$ for all $k$. Then

$$\tilde{P}_t = \lim_{\text{sot}} \pi(P_{k,m_k}) = \lim_{\text{sot}} \pi(U)\pi(P_{k,1})\pi(U)^* = \pi(U)\tilde{P}\pi(U)^* \neq 0. \qquad \square$$

**14.7.1. The Fundamental Group.** The name **fundamental group** is usually saved to describe the group of equivalence classes of loops of a topological space under homotopy. Here we use the term—because of historical reasons—with a completely independent meaning.

Given a $\mathrm{II}_1$-factor $\mathcal{M}$, with $\tau$ its unique normalized trace, and a nonzero projection $P \in \mathcal{M}$, we know from Proposition 14.1.1 that $P\mathcal{M}P$ is a factor. It is finite, because $P$ is finite, and it is infinite-dimensional because it has no abelian projections; it then has to be a $\mathrm{II}_1$ factor. It is a natural question to ask, and it occurred already to Murray and von Neumann, when is $P\mathcal{M}P \simeq \mathcal{M}$? When $\mathcal{M}$ is the hyperfinite $\mathrm{II}_1$-factor, Murray–von Neumann proved that it is unique, and that all its subfactors are hyperfinite. Hence $P\mathcal{M}P \simeq \mathcal{M}$ for all $P$. For other factors, the question is subtle.

We know from Exercise 14.2.15 and Theorem 14.5.10 that in an arbitrary $\mathrm{II}_1$-factor we have $P\mathcal{M}P \simeq Q\mathcal{M}Q$ when $\tau(P) = \tau(Q)$. So given $t \in (0, 1]$ we have a (class of) subfactor(s) $\mathcal{M}_t \subset \mathcal{M}$, where $\mathcal{M}_t = P\mathcal{M}P$ for some projection $P \in \mathcal{M}$ with $\tau(P) = t$.

Here is a small result that suggests this stuff is interesting.

***14.7.7. Lemma.***

*Let $s, t \in (0, 1]$ such that $\mathcal{M}_s \simeq \mathcal{M}_t \simeq \mathcal{M}$. Then $\mathcal{M}_{st} \simeq \mathcal{M}$.*

*Proof.* Let $P \in \mathcal{P}(\mathcal{M})$ with $\tau(P) = t$. Since $P\mathcal{M}P \simeq \mathcal{M}$ and $\mathcal{M}_s \simeq \mathcal{M}$, there is a projection $Q \in P\mathcal{M}P$ with $(P\mathcal{M}P)_s \simeq Q\mathcal{M}Q$. So $\tau_P(Q) = s$, where $\tau_P$ indicates the unique normalized trace on $P\mathcal{M}P$. This trace is necessarily $\tau_P(T) = \tau(T)/\tau(P)$ (because this gives a normal faithful trace on $P\mathcal{M}P$ and the uniqueness). Hence

$$s = \tau_P(Q) = \frac{\tau(Q)}{\tau(P)} = \frac{1}{t}\,\tau(Q).$$

Then $\tau(Q) = st$ and so $M_{st} \simeq \mathcal{M}$ since

$$M_{st} \simeq Q\mathcal{M}Q \simeq P\mathcal{M}P \simeq \mathcal{M}.$$ □

Consider the case where $\mathcal{M}_{1/2} \simeq \mathcal{M}$ (this is true for instance for the hyperfinite $\mathrm{II}_1$-factor). This means that there exists a projection $P \in \mathcal{M}$ with $\tau(P) = 1/2$ and $P\mathcal{M}P \simeq \mathcal{M}$. Fix a $*$-isomorphism $\gamma : \mathcal{M} \to P\mathcal{M}P$. Since $\tau(I_\mathcal{M} - P) = 1/2$, we also have $(I_\mathcal{M} - P)\mathcal{M}(I_\mathcal{M} - P) \simeq \mathcal{M}$. Let $V \in \mathcal{M}$ be a partial isometry with $V^*V = P$ and $VV^* = I_\mathcal{M} - P$. Then $\{P, V, V^*, I_\mathcal{M} - P\}$ behave as $2 \times 2$ matrix units. Let $\theta : \mathcal{M} \to M_2(\mathcal{M})$ be the map

$$\theta(T) = \begin{bmatrix} \gamma^{-1}(PTP) & \gamma^{-1}(PT(I_\mathcal{M} - P)V) \\ \gamma^{-1}(V^*(I_\mathcal{M} - P)TP) & \gamma^{-1}(V^*(I_\mathcal{M} - P)T(I_\mathcal{M} - P)V) \end{bmatrix}. \tag{14.24}$$

A straightforward computation shows that $\theta$ is an injective $*$-homomorphism (Exercise 14.7.3). And given $\tilde{T} \in M_2(\mathcal{M})$ we can form

$$T = \gamma(T_{11}) + \gamma(T_{12})V^* + V\gamma(T_{21}) + V\gamma(T_{22})V^* \in \mathcal{M}.$$

Then, since $V^*P = PV = 0$,

$$\gamma^{-1}(PTP) = \gamma^{-1}(\gamma(T_{11})) = T_{11}.$$

Similarly,

$$\begin{aligned}\gamma^{-1}(PT(I_{\mathcal{M}} - P)V) &= \gamma^{-1}(\gamma(T_{12})) = T_{12},\\ \gamma^{-1}(V^*(I_{\mathcal{M}} - P)TP) &= \gamma^{-1}(\gamma(T_{21})) = T_{21},\\ \gamma^{-1}(V^*(I_{\mathcal{M}} - P)T(I_{\mathcal{M}} - P)V) &= \gamma^{-1}(\gamma(T_{22})) = T_{22}.\end{aligned}$$

Hence $\theta$ is surjective and we have a natural isomorphism $\mathcal{M} \simeq M_2(\mathcal{M})$. With the same ideas, if $\mathcal{M}_{1/n} \simeq \mathcal{M}$ then $\mathcal{M} \simeq M_n(\mathcal{M})$. And these implications also go back: if $\mathcal{M} \simeq M_n(\mathcal{M})$, then $\mathcal{M} \simeq E_{11}M_n(\mathcal{M})E_{11}$, and $E_{11}$ has trace $1/n$ in $M_n(\mathcal{M})$, so $\mathcal{M}_{1/n} \simeq \mathcal{M}$.

For $t \in (0, \infty)$ we can consider $\mathcal{M}_t$ to be $M_n(\mathcal{M})_{t/n}$, where $n \in \mathbb{N}$ satisfies $n > t$. This is coherent, in the following sense:

**14.7.8. Lemma.**

*Let $\mathcal{M}$ be a $\mathrm{II}_1$-factor, $t \in (0, \infty)$, and $n, k \in \mathbb{N}$ with $t < n < k$. Fix projections $P \in M_n(\mathcal{M})$ and $Q \in M_k(\mathcal{M})$ such that $\tau_n(P) = t/n$ and $\tau_k(Q) = t/k$. Then*

$$PM_n(\mathcal{M})P \simeq QM_k(\mathcal{M})Q.$$

*Proof.* We can embed $M_n(\mathcal{M}) \subset M_k(\mathcal{M})$ as the top-left corner, and consider the projection $\tilde{P} = P \oplus 0$ (that is, put $P$ on the top-left corner and fill the rest of the matrix with zeros). The trace of $\tilde{P}$ is

$$\tau_k(\tilde{P}) = \frac{1}{k}\sum_{j=1}^{k} \tau(\tilde{P}) = \frac{1}{k}\sum_{j=1}^{n} \tau(P) = \frac{n}{k}\,\tau_n(P) = \frac{t}{k}.$$

Therefore, because of Exercise 14.2.15,

$$PM_n(\mathcal{M})P \simeq \tilde{P}M_k(\mathcal{M})\tilde{P} \simeq QM_k(\mathcal{M})Q. \qquad \square$$

With Lemma 14.7.8 at hand, we can define:

**14.7.9. Definition.**

*Let $\mathcal{M}$ be a $\mathrm{II}_1$-factor. The set*

$$\mathcal{F}(\mathcal{M}) = \{t \in (0, \infty) : \ \mathcal{M}_t \simeq \mathcal{M}\}$$

*is called the **Fundamental Group** of $\mathcal{M}$.*

And the name is justified:

***14.7.10. Proposition.***

*The set $\mathcal{F}(\mathcal{M})$, with the operation multiplication, is a group.*

*Proof.* Let $s, t \in \mathcal{F}(\mathcal{M})$. Fix $n \in \mathbb{N}$ with $n \geq \max\{s, t\}$. By definition, we have

$$M_n(\mathcal{M})_{s/n} \simeq M_n(\mathcal{M})_{t/n} \simeq \mathcal{M}.$$

This means that there exists a projection $P \in M_n(\mathcal{M})$ with $\tau_n(P) = s/n$ and $PM_n(\mathcal{M})P \simeq \mathcal{M}$. Since $t \in \mathcal{F}(\mathcal{M})$, we have that $(PM_n(\mathcal{M})P)_t \simeq \mathcal{M}$. So there exists a projection $Q \in M_n(PM_n(\mathcal{M})P)$ with $\tilde{\tau}(Q) = t/n$ and

$$QM_n(PM_n(\mathcal{M})P)Q \simeq \mathcal{M},$$

where $\tilde{\tau}$ is the normalized trace on $M_n(PM_n(\mathcal{M})P)$. The normalized trace on $PM_n(\mathcal{M})P$ is $\tau_{n,P}$ given by $\tau_{n,P}(X) = \tau_n(X)/\tau_n(P)$. Then (using Proposition 14.2.22 to subdivide $Q$)

$$\frac{t}{n} = \tilde{\tau}(Q) = \frac{1}{n}\sum_{j=1}^{n} \tau_{n,P}(Q_{jj}) = \frac{1}{n}\sum_{j=1}^{n} \frac{\tau_n(Q_{jj})}{\tau_n(P)} = \frac{1}{s}\sum_{j=1}^{n} \tau_n(Q_{jj}).$$

So

$$\frac{st}{n^2} = \frac{1}{n}\sum_{j=1}^{n} \tau_n(Q_{jj}) = \frac{1}{n^2}\sum_{j=1}^{n} \tau(Q_{jj}) = \tau_{n^2}(Q).$$

The last equality is justified by thinking of $Q \in M_{n^2}(\mathcal{M})$ by putting $Q$ in the upper left corner. Then $QM_{n^2}(\mathcal{M})Q = QM_n(PM_n(\mathcal{M})P)Q \simeq \mathcal{M}$ (we leave the proof of the equality as Exercise 14.7.4), and therefore $st \in \mathcal{F}(\mathcal{M})$ and $\mathcal{F}(\mathcal{M})$ is multiplicative.

The previous paragraph also shows that $(\mathcal{M}_s)_t \simeq \mathcal{M}_{st}$. Given $t \in \mathcal{F}(\mathcal{M})$, we have $\mathcal{M} \simeq \mathcal{M}_t$. Then

$$\mathcal{M}_{1/t} \simeq (\mathcal{M}_t)_{1/t} \simeq \mathcal{M}_{t\,1/t} \simeq \mathcal{M}.$$

Then $1/t \in \mathcal{F}(\mathcal{M})$ and thus $\mathcal{F}(\mathcal{M})$ is a multiplicative subgroup of $(0, \infty)$. □

As mentioned above, when $\mathcal{R}$ is the hyperfinite $\mathrm{II}_1$-factor every compression $\mathcal{R}_t$ is again a $\mathrm{II}_1$-factor, and Murray–von Neumann proved that every subfactor of $\mathcal{R}$ is again hyperfinite, and that there is a unique hyperfinite $\mathrm{II}_1$-factor. It follows that $\mathcal{R}_t \simeq \mathcal{R}$ for all $t \in (0, 1]$ and this implies that $\mathcal{F}(\mathcal{R}) = (0, \infty)$. This raises the question of what is $\mathcal{F}(\mathcal{M})$ for other $\mathrm{II}_1$-factors. For 40 years after Murray-von Neumann there was zero progress on this. In 1980 Connes was able to show that if $G$ is an icc countable discrete group that satisfies Kazhdan's property T, then $\mathcal{F}(L(G))$ is countable. This, together with [**FM77a**, **FM77b**], sparked lots of research on $\mathrm{II}_1$-factors. Voiculescu's Free Probability, that exploded in the early 1990s with his own work and that of Dykema and Radulescu, produced the so called free group factors, which are a continuous $L(\mathbb{F}_t)$ of $\mathrm{II}_1$-factors; Dykema and Radulescu [**Dyk94**, **Ř94**] showed independently that either these factors are all isomorphic (which happens precisely if $\mathcal{F}(L(\mathbb{F}_t)) = (0, \infty)$ for all $t$), or their are all pairwise non-isomorphic (which happens precisely if $\mathcal{F}(L(\mathbb{F}_t)) = \{1\}$ for all $r$). This would have settled the isomorphism problem for the factors $L(\mathbb{F}_n)$, but surprisingly the question remains open to this day. It was more than 20 years after Connes paper, that Gabouriau's remarkable work on Borel equivalence relations [**Gab02**] inspired to Popa's foundational [**Pop06a**], where among other byproducts the first example of a $\mathrm{II}_1$-factor with $\mathcal{F}(\mathcal{M}) = \{1\}$ was exhibited. One such example, as has been mentioned, is $L(\mathbb{Z}_2 \,\overline{\rtimes}\, SL_2(\mathbb{Z}))$. Many results, by Popa, his students, and others, followed. In [**Pop06b**] Popa showed that given any countable subgroup $G \subset (0, \infty)$ there exists an action of $SL_2(\mathbb{Z})$ on the hyperfinite $\mathrm{II}_1$-factor $\mathcal{R}$ such that $\mathcal{F}(\mathcal{R} \rtimes SL_2(\mathbb{Z})) = G$. Decades later, Popa's rigidity/deformation theory and the study of $\mathrm{II}_1$-factors of the form $L^\infty(X) \,\overline{\rtimes} G$ where the action of $G$ on $L^\infty(X)$ is induced by a measure-preserving action of $G$ on $X$, is an active area of research.

**14.7.2. Exercises.**

(14.7.1) Where in the proof of Proposition 14.7.2 is the norm-closedness of $\mathcal{J}$ used?

(14.7.2) Give an alternative proof of Proposition 14.7.2 by using Dixmier's Property.

(14.7.3) Show that $\theta$ in (14.24) is a $*$-homomorphism.

(14.7.4) Let $\mathcal{M}$ be a $\mathrm{II}_1$-factor, $n \in \mathbb{N}$,

$$P \in \mathcal{P}(M_n(\mathcal{M})), \quad \text{and} \quad Q \in M_n(PM_n(\mathcal{M})P).$$

Show that $QM_{n^2}(\mathcal{M})Q = QM_n(PM_n(\mathcal{M})P)Q$.

APPENDIX A

# The Determinant

The notion of determinant is key in linear algebra. While its use is pretty basic and well-known, it is not that common to find a detailed account of its existence and properties, which we tackle below. While only the usual field $\mathbb{R}$ and $\mathbb{C}$ are needed in this book, nothing changes below if one uses an arbitrary field $\mathbb{F}$, so that is the choice we use.

## A.1. Preliminaries on Permutations

**A.1.1. Definition.**

*For $n \in \mathbb{N}$ the **permutation group** $\mathbb{S}_n$ is the set of bijections $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$, with the operation composition.*

*A **transposition** is a permutation $\sigma \in \mathbb{S}_n$ such that there exist $k, j \in \{1,\ldots,n\}$ with $\sigma(k) = j$, $\sigma(j) = k$, and $\sigma(r) = r$ for all $r \neq k, j$. In such case we denote $\sigma$ by $(k\ j)$.*

*A **cycle** is a permutation $\sigma \in \mathbb{S}_n$ such that there exist $a_1,\ldots,a_m \in \{1,\ldots,n\}$ such that $\sigma(a_k) = a_{k+1}$, $k = 1,\ldots,m-1$, and $\sigma(a_m) = a_1$. We use the notation $\sigma = (a_1\ a_2\ \cdots\ a_m)$.*

As composition of bijections is a bijection and bijections are invertible, $\mathbb{S}_n$ is a group under composition. This prompts us to use the word "product" to refer to composition in this context.

***A.1.2. Proposition.***

*Every permutation can be written in a unique way, up to order, as a product of disjoint cycles.*

*Proof.* We prove the existence by induction. When $n = 2$, the only two permutations are $(1)(2)$ and $(1\ 2)$. Now let $\sigma \in \mathbb{S}_n$. Let $a_1 = 1$ and $\{a_1, \ldots, a_m\} = \{\sigma^{k-1}(1) : k \in \mathbb{N}\}$. The set is finite because the range of $\sigma$ is finite. We choose $m$ to be the least positive integer such that there exists $0 \leq p < m$ such that $\sigma^m(1) = \sigma^p(1)$. If $p > 0$, as $\sigma$ is invertible, we get $\sigma^{m-1}(1) = \sigma^{p-1}(1)$, contradicting the minimality. Hence $p = 0$ and $\sigma^m(1) = 1$. Thus we can take $a_j = \sigma^{j-1}(1)$, and then $a_{j+1} = \sigma(a_j)$, $\sigma(a_m) = 1$. That is,

$$\sigma = (a_1\ a_2\ \cdots\ a_m).$$

By inductive hypothesis, $\sigma$ restricted to the set $\{1, \ldots, n\} \setminus \{a_1, \ldots, a_m\}$ is a product of disjoint cycles, and hence $\sigma$ is a product of cycles.

As for the uniqueness, the construction in the previous paragraph shows that each cycle has to necessarily exist in the presentation of $\sigma$. So the only possibility one needs to discard is that a cycle be written as a product of smaller cycles; but this is impossible because it would contradict the minimality of $m$ above. $\square$

***A.1.3. Corollary.***

*Every permutation can be written as a product of transpositions.*

*Proof.* In light of Proposition A.1.2, it is enough to show that a cycle is a product of transpositions. And this is simply, with the assumption that $a_1 < a_2 < \cdots < a_m$,

$$(a_1\ a_2\ \cdots\ a_m) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{m-1}\ a_m). \qquad \square$$

The presentation of $\sigma$ as a product of transpositions is not unique, but we will momentarily see that the parity of the number of transpositions used to write $\sigma$ is always the same.

**A.1.4. Definition.**

*Given $\sigma \in \mathbb{S}_n$, the* sign *of $\sigma$ is the number $\operatorname{sgn}\sigma \in \{-1,1\}$, determined in the following way. Let*

$$p(x_1,\dots,x_n) = \prod_{1\le j<k\le n} (x_j - x_k).$$

*Then*

$$\operatorname{sgn}\sigma = \frac{p(x_{\sigma(1)},\dots,x_{\sigma(n)})}{p(x_1,\dots,x_n)}.$$

For instance if $\sigma = (1\ 2) \in \mathbb{S}_3$, then $p(x_1,x_2,x_3) = (x_1-x_2)(x_1-x_3)(x_2-x_3)$, and

$$\operatorname{sgn}\sigma = \frac{(x_2-x_1)(x_2-x_3)(x_1-x_3)}{(x_1-x_2)(x_1-x_3)(x_2-x_3)} = -1.$$

***A.1.5. Proposition.***

*Let $\sigma,\eta \in \mathbb{S}_n$. Then*

*(i)* $\operatorname{sgn}\sigma\eta = (\operatorname{sgn}\sigma)(\operatorname{sgn}\eta)$*;*

*(ii)*

$$\operatorname{sgn}\sigma = \begin{cases} 1, & \sigma \text{ can be written as the product of an even number of transpositions} \\ -1, & \sigma \text{ can be written as the product of an odd number of transpositions} \end{cases}$$

*(iii)*

$$\operatorname{sgn}\sigma = \begin{cases} 1, & |\{(r,s):\ r<s,\ \sigma(r)>\sigma(s)\}| \text{ is even} \\ -1, & |\{(r,s):\ r<s,\ \sigma(r)>\sigma(s)\}| \text{ is odd} \end{cases}$$

*Proof.* It's remarkable how easily we can get the three properties out of the definition of $\operatorname{sgn}\sigma$.

(i) We have

$$\begin{aligned}\operatorname{sgn}\sigma\eta &= \frac{p(x_{\sigma\eta(1)},\dots,x_{\sigma\eta(n)})}{p(x_1,\dots,x_n)} \\ &= \frac{p(x_{\sigma\eta(1)},\dots,x_{\sigma\eta(n)})}{p(x_{\eta(1)},\dots,x_{\eta(n)})}\,\frac{p(x_{\eta(1)},\dots,x_{\eta(n)})}{p(x_1,\dots,x_n)} \\ &= \operatorname{sgn}\sigma\ \operatorname{sgn}\eta.\end{aligned}$$

(ii) We know that $\operatorname{sgn}(1\ 2) = -1$. For any other transposition,

$$\begin{aligned}\operatorname{sgn}(r\ s) &= \operatorname{sgn}(1\ r)(2\ s)(1\ 2)(2\ s)(1\ r)\\ &= (-1)[\operatorname{sgn}(1\ r)]^2[\operatorname{sgn}(2\ s)]^2 = -1.\end{aligned}$$

That is, $\operatorname{sgn} t = -1$ for any transposition $t$. So, if $\sigma = t_1 t_2 \cdots t_m$ is a product of $m$ transpositions, then $\operatorname{sgn}\sigma = (-1)^m$.

(iii) Each pair $(r, s)$ with $r < s$ and $\sigma(r) > \sigma(s)$, is a pair that in the definition of $\operatorname{sgn}\sigma$ will contribute

$$\frac{(x_{\sigma(r)} - x_{\sigma(s)})}{(x_s - x_r)} < 0,$$

while the rest of the pairs $r < s$, where $\sigma(r) < \sigma(s)$, contribute with

$$\frac{(x_{\sigma(r)} - x_{\sigma(s)})}{(x_r - x_s)} > 0.$$

Thus if $c = |\{(r, s) :\ r < s,\ \sigma(r) > \sigma(s)\}|$, then $\operatorname{sgn}\sigma = (-1)^c$. □

## A.2. Preliminaries on Multilinear Maps

**A.2.1. Definition.**

*Let $V$ be a vector space over the field $\mathbb{F}$. A function $f : V^m \to \mathbb{F}$ is*

- ***multilinear**, if*

$$\begin{aligned}f(x_1, \ldots, x_{j-1}, &ax + by, x_{j+1}, \ldots, x_m)\\ &= af(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_m)\\ &\quad + bf(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_m)\end{aligned}$$

  *for all $a, b \in \mathbb{F}$, all $j \in \{1, \ldots, n\}$, and all $x, y, x_1, \ldots, x_m \in V$;*
- ***skew-symmetric** if $f(x_{\sigma(1)}, \ldots, x_{\sigma(m)}) = \operatorname{sgn}\sigma\, f(x_1, \ldots, x_m)$ for all $x_1, \ldots, x_m \in V$ and all $\sigma \in \mathbb{S}_n$;*
- ***alternating** if $f(x_1, \ldots, x_m) = 0$ whenever $x_1, \ldots, x_m \in V$ and there exist $k \neq j$ with $x_k = x_j$.*

***A.2.2. Proposition.***

*A multilinear alternating function $f : V^m \to \mathbb{F}$ is skew-symmetric. When $\mathbb{F}$ does not have characteristic 2, the converse holds, that is if $f$ is alternating then it is skew-symmetric.*

*Proof.* Suppose that $f$ is alternating, and $\sigma = (k\ j)$. Then, using that $f$ is multilinear and alternating (so that two terms in the expansion below disappear as they are zero),

$$\begin{aligned} 0 &= f(x_1, \ldots, x_{j-1}, x_j + x_k, x_{j+1}, \ldots, x_{k-1}, x_j + x_k, x_{k+1}, \ldots, x_m) \\ &= f(x_1, \ldots, x_m) + f(x_1, \ldots, x_{j-1}, x_k, x_{j+1}, \ldots, x_{k-1}, x_j, x_{k+1}, \ldots, x_m) \\ &= f(x_1, \ldots, x_m) + f(x_{\sigma(1)}, \ldots, x_{\sigma(m)}). \end{aligned}$$

Thus

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(m)}) = -f(x_1, \ldots, x_m) = \operatorname{sgn}\sigma\, f(x_1, \ldots, x_m). \tag{A.1}$$

For arbitrary $\sigma \in \mathbb{S}_n$, it can be written as a composition $\sigma_1 \cdots \sigma_r$ of transpositions by Corollary A.1.3. As $\operatorname{sgn}\sigma = \prod_j \operatorname{sgn}\sigma_j$, it follows that $f$ is skew-symmetric by repeated application of (A.1).

Now assume that the characteristic of $\mathbb{F}$ is not 2 and $f$ is skew-symmetric. If $x_k = x_j$ for $k \neq j$, let $\sigma = (k\ j)$. Since $\operatorname{sgn}\sigma = -1$ and the list $x_1, \ldots, x_m$ does not change if we permute $x_k$ and $x_j$, we get that

$$f(x_1, \ldots, x_m) = \operatorname{sgn}\sigma\, f(x_1, \ldots, x_m) = -f(x_1, \ldots, x_m).$$

Hence $f(x_1, \ldots, x_m) = 0$ and $f$ is alternating. □

***A.2.3. Theorem.***

*Let $V$ be a vector space and $e_1, \ldots, e_n$ a basis for $V$. Then there exists a unique alternating multilinear function $f : V^n \to \mathbb{F}$ such that $f(e_1, \ldots, e_n) = 1$. Explicitly,*

$$f\Big(\sum_{j_1=1}^{n} a_{1j}e_j, \ldots, \sum_{j_n=1}^{n} a_{nj}e_j\Big) = \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

*Proof. (Uniqueness).* Let $v_1, \ldots, v_n \in V$. We can write

$$v_k = \sum_{j=1}^{n} a_{kj}e_j, \qquad k = 1, \ldots, n$$

for appropriate (and unique) coefficients $a_{11}, \ldots, a_{nn} \in \mathbb{F}$. Then, since $f$ is alternating and hence all terms where $j_k = j_h$ for some $h, k$ are zero, and using also that $f$ is skew-symmetric (Proposition A.2.2),

$$\begin{aligned}
f(v_1,\dots,v_n) &= f\Big(\sum_{j_1=1}^{n} a_{1j}e_j,\dots,\sum_{j_n=1}^{n} a_{nj}e_j\Big)\\
&= \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n} a_{1,j_1}\cdots a_{n,j_n}\, f(e_{j_1},\dots,e_{j_n})\\
&= \sum_{\sigma\in\mathbb{S}_n} a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}\, f(e_{\sigma(1)},\dots,e_{\sigma(n)})\\
&= \sum_{\sigma\in\mathbb{S}_n} a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}\, \operatorname{sgn}\sigma\, f(e_1,\dots,e_n)\\
&= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.
\end{aligned}$$

This last expression does not depend on $f$, so the uniqueness is proven. It also tells us how to define $f$!

*(Existence)* We define

$$f\Big(\sum_{j_1=1}^{n} a_{1j}e_j,\dots,\sum_{j_n=1}^{n} a_{nj}e_j\Big) = \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.$$

We need to show that $f$ is an alternating multilinear form on $V$. The multilinearity is straightforward, as the product $a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}$ contains exactly one factor from each variable in the domain of $f$. We also have $f(e_1,\dots,e_n)=1$, since in such case we have $a_{kk}=1$ for all $k$ and $a_{kj}=0$ for $k\neq j$.

Now suppose that $\sum_{j=1}^{n} a_{kj}e_j = \sum_{j=1}^{n} a_{hj}e_j$ for some $h\neq k$. That is, $a_{kj}=a_{hj}$ for some $h\neq k$ and all $j$. Let $\rho=(k\ h)$. For any $\eta\in\mathbb{S}_n$, $\operatorname{sgn}(\rho\eta) = -\operatorname{sgn}(\eta)$. We also have

$$a_{1,\eta(1)}\cdots a_{n,\eta(n)} = a_{1,\rho\eta(1)}\cdots a_{n,\rho\eta(n)},$$

since

$$a_{h,\eta(h)}a_{k,\eta(k)} = a_{k,\eta(h)}a_{h,\eta(k)} = a_{h,\eta\rho(h)}a_{k,\eta\rho(k)}$$

and $a_{r,\eta\rho(r)} = a_{r,\eta(r)}$ for all $r\notin\{h,k\}$. Then

$$\begin{aligned}
f\Big(\sum_{j_1=1}^{n} a_{1j}e_j,\dots,\sum_{j_n=1}^{n} a_{nj}e_j\Big) &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}\\
&= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma\rho(1)}\cdots a_{n,\sigma\rho(n)} = -\sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\rho\, a_{1,\sigma\rho(1)}\cdots a_{n,\sigma\rho(n)}\\
&= -\sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} = -f\Big(\sum_{j_1=1}^{n} a_{1j}e_j,\dots,\sum_{j_n=1}^{n} a_{nj}e_j\Big),
\end{aligned}$$

and so $f$ is alternating. $\square$

## A.3. The Determinant

**A.3.1. Definition.**

*The **determinant** of $A \in M_n(\mathbb{F})$ is the number*

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\, M_{1j}^{A},$$

*where the **minor** $M_{kj}^{A}$ is recursively defined as the determinant of the matrix obtained from $A$ by removing the $k^{\text{th}}$ row and the $j^{\text{th}}$ column, and $M_{11}^{A} = 1$ when $n = 1$.*

***A.3.2. Proposition.***

*Let $A \in M_n(\mathbb{F})$. Then*

$$\det A = \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn} \sigma\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \tag{A.2}$$

*Proof.* We proceed by induction on $n$. When $n = 1$, there is nothing to prove. So we assume that (A.2) holds for matrices of size less than $n$. Let $B^{(j)} = A(1, j)$, the matrix obtained by removing row 1 and column $j$ from $A$ . We have

$$b_{r,s}^{(j)} = \begin{cases} a_{r+1,s}, & s < j \\ a_{r+1,s+1}, & s \geq j \end{cases}$$

For fixed $j \in \{1, \ldots, n\}$ and $\nu \in \mathbb{S}_{n-1}$, we define $\sigma_{j,\nu} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ by

$$\sigma_{j,\nu}(k) = \begin{cases} j, & k = 1 \\ \nu(k-1), & k \geq 2,\ \nu(k-1) < j \\ \nu(k-1) + 1, & k \geq 2,\ \nu(k-1) \geq j \end{cases} \tag{A.3}$$

A straightforward look shows that $\sigma_{j,\nu}$ is injective (Exercise A.3.1), and hence $\sigma_{j,\nu} \in \mathbb{S}_n$. We also have

$$\operatorname{sgn} \sigma_{j,\nu} = (-1)^{j-1} \operatorname{sgn} \nu$$

(Exercise A.3.2). Note also that

$$\{\sigma_{j,\nu} :\ j \in \{1, \ldots, n\},\ \nu \in \mathbb{S}_{n-1}\} = \mathbb{S}_n$$

(Exercise A.3.3).

Now we can calculate, using the inductive hypothesis on $M_{1j}^A$,

$$\begin{aligned}\det A &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\, M_{1j}^A \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \sum_{\nu\in\mathbb{S}_{n-1}} \operatorname{sgn}\nu\, b^{(j)}_{1,\nu(2)} \cdots b^{(j)}_{n-1,\nu(n)} \\ &= \sum_{j=1}^{n} \sum_{\nu\in\mathbb{S}_{n-1}} (-1)^{1+j} \operatorname{sgn}\nu\, a_{1,\sigma_{j,\nu}(1)} \cdots a_{n,\sigma_{j,\nu}(n)} \\ &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}\sigma\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \qquad \square\end{aligned}$$

### *A.3.3. Theorem.*

*The function* $\det : M_n(\mathbb{F}) \to \mathbb{F}$ *has the following properties:*

*(i)* *if $B$ is obtained from $A$ by multiplying a row by $\alpha$, then* $\det B = \alpha \det A$*;*

*(ii)* *if we have $A, B, C$ with $A_{rj} = B_{rj} + C_{rj}$ for all $j$, and $A_{kj} = B_{kj} = C_{kj}$ when $k \neq r$ and for all $j$, then* $\det A = \det B + \det C$*;*

*(iii)* *if $A$ has two equal rows, then* $\det A = 0$*;*

*(iv)* *if $B$ is obtained from $A$ by swapping two rows, then* $\det B = -\det A$*;*

*(v)* *if $B_{rj} = A_{rj} + \alpha A_{sj}$, and $B_{kj} = A_{kj}$ when $k \neq r$, then* $\det B = \det A$;

*(vi)* *for any elementary matrix $E$,*
$$\det EA = \det E\, \det A.$$

*(vii)* $\det AB = \det A \det B$*.*

*(viii)* *$A$ is invertible if and only if* $\det A \neq 0$*; in that case,* $\det A^{-1} = \frac{1}{\det A}$*;*

*(ix)* *if $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $A$, counting multiplicities, then* $\det A = \lambda_1 \cdots \lambda_n$*;*

*(x)* $\det A^\top = \det A$*;*

*(xi) for any* $k = 1, \dots, n$

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} M_{kj}^A; \tag{A.4}$$

*(xii) for any* $j = 1, \dots, n$

$$\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} M_{kj}^A; \tag{A.5}$$

*Proof.* Properties (i)–(v) are automatic when det is seen via Theorem A.2.3 as the unique normalized alternating multilinear map on the rows of $A$, seen as vectors in the canonical basis.

(vi) The elementary matrix corresponding to multiplying row $k$ by $\alpha$ has rows

$$e_1, \dots, e_{k-1}, \alpha e_k, e_{k+1}, \dots, e_n.$$

Then, using det to also denote the multilinear map,

$$\det E = \det(e_1, \dots, e_{k-1}, \alpha e_k, e_{k+1}, \dots, e_n) = \alpha \det(e_1, \dots, e_n) = \alpha,$$

and

$$\det EA = \alpha \det A = \det E \det A.$$

The elementary matrix $E$ corresponding to permuting rows $k$ and $h$ has rows

$$e_1, \dots, e_{k-1}, \alpha e_h, e_{k+1}, \dots, e_{h-1}, e_k, e_{h+1}, \dots, e_n.$$

Then

$$\begin{aligned}\det E &= \det(e_1, \dots, e_{k-1}, \alpha e_h, e_{k+1}, \dots, e_{h-1}, e_k, e_{h+1}, \dots, e_n)\\ &= -\det(e_1, \dots, e_n) = -1,\end{aligned}$$

and hence

$$\det EA = -\det A = \det E \det A.$$

Finally, the elementary matrix corresponding to adding $\alpha$ times row $h$ to row $k$ has rows

$$e_1, \dots, e_{k-1}, e_k + \alpha e_h, e_{k+1}, \dots, e_n.$$

Then

$$\begin{aligned}\det E &= \det(e_1, \dots, e_{k-1}, e_k + \alpha e_h, e_{k+1}, \dots, e_n)\\ &= \det(e_1, \dots, e_n) + \alpha \det(e_1, \dots, e_{k-1}, e_h, e_{k+1}, \dots, e_n)\\ &= \det(e_1, \dots, e_n) = 1,\end{aligned}$$

since the second term above has $e_h$ twice. And so

$$\det EA = \det A = \det E \det A.$$

(vii) If $B$ is invertible, there exist elementary matrices $E_1, \ldots, E_m$ such that $B = E_1 \cdots E_m$. Then

$$\det BA = \det E_1 \cdots E_m A = \det E_1 \cdots \det E_m \det A$$
$$= \det(E_1 \cdots E_m) \det A = \det B \det A.$$

And if $B$ is not invertible, there exists an invertible matrix such that $SB$ has a row of all zeros (where $SB$ is the reduced echelon form of $B$). Since $SS^{-1} = I_n$, $\det S \; \det S^{-1} = 1$, and so $\det S \neq 0$. Then

$$0 = \det SB = \det S \; \det B,$$

and thus $\det B = 0$. As $BA$ cannot be invertible if $B$ isn't (if $BA$ is invertible then $B$ is injective, and hence surjective as the codomain has the same dimension as the domain),

$$0 = \det BA = \det B \; \det A.$$

Hence $\det BA = \det B \det A$ holds for all pairs of $A, B \in M_n(\mathbb{C})$.

(viii) If $A$ is invertible then $\det A \; \det A^{-1} = 1$, so

$$\det A^{-1} = \frac{1}{\det A}.$$

(ix) If $J$ is the Jordan form of $A$ (which always exists over the algebraic closure of $\mathbb{F}$), there exists $S$ invertible such that $A = SJS^{-1}$. Since $J$ is triangular with diagonal $\lambda_1, \ldots, \lambda_n$, the only diagonal product that is nonzero is $\lambda_1 \cdots \lambda_n$, and so $\det J = \lambda_1 \cdots \lambda_n$. Thus

$$\det A = \det SJS^{-1} = \det S \; \det J \; \det S^{-1} = \det J = \lambda_1 \cdots \lambda_n.$$

(x) As $A^\top$ has the same eigenvalues as $A$,

$$\det A^\top = \lambda_1 \cdots \lambda_n = \det A.$$

(xi) By definition,

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \, M^A_{1j}.$$

If $B$ is the matrix obtained by exchanging rows 1 and 2 in $A$, then

$$\det A = -\det B = -\sum_{j=1}^{n} (-1)^{1+j} b_{1j} \, M^B_{1j} = \sum_{j=1}^{n} (-1)^{2+j} a_{2j} \, M^A_{2j}.$$

Inductively, if we assume that

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \, M^A_{kj}$$

and $B$ is the matrix obtained by exchanging rows $k$ and $k+1$ on $A$,

$$\det A = -\det B = -\sum_{j=1}^{n}(-1)^{k+j} b_{kj}\, M^B_{kj} = \sum_{j=1}^{n}(-1)^{(k+1)+j} a_{k+1,j}\, M^A_{k+1,j}.$$

This shows, by induction, that (A.4) holds.

(xii) We have

$$\det A = \det A^\top = \sum_{k=1}^{n}(-1)^{j+k}(A^\top)_{jk}\, M^{A^\top}_{jk} = \sum_{k=1}^{n}(-1)^{k+j} a_{kj}\, M^A_{kj}. \qquad \square$$

**A.3.1. Exercises.**

(A.3.1) Given $j \in \{1,\dots,n\}$ and $\nu \in \mathbb{S}_{n-1}$, let $\sigma_{j,\nu}$ as in (A.3). Show that $\sigma_{j,\nu} \in \mathbb{S}_n$.

(A.3.2) Given $j \in \{1,\dots,n\}$ and $\nu \in \mathbb{S}_{n-1}$, let $\sigma_{j,\nu}$ as in (A.3). Show that $\operatorname{sgn}\sigma_{j,\nu} = (-1)^{j-1}\operatorname{sgn}\nu$.

(A.3.3) With the notation of (A.3), show that $\mathbb{S}_n = \{\sigma_{j,\nu} :\ j \in \{1,\dots,n\},\ \nu \in \mathbb{S}_{n-1}\}$.

APPENDIX

# B

# Getting to Know Majorization

Majorization is a notion in Matrix Analysis that somehow describes how vectors are weighted. It is intimately related with certain kind of inequalities that appear when considering eigenvalues and singular values of matrices. We will develop here the parts of the theory needed to eventually obtain certain inequalities that will be needed soon. But majorization goes well beyond that, and it has all kinds of applications, from Matrix Analysis to Economics to Quantum Information to Signal Processing and more. We refer to [**MOA11**] as a major reference and point of entry to the topic.

## B.1. Preliminaries on Majorization

For deeper details than those covered here we recommend [**HJ94**, **Bha97**].

**B.1.1. Definition.**

*For $x \in \mathbb{R}^n$, we denote by $x^{\downarrow} \in \mathbb{R}^n$ the element obtained from $x$ by reordering it in non-increasing order: $x^{\downarrow}_1 \geq x^{\downarrow}_2 \geq \cdots \geq x^{\downarrow}_n$. Similarly, $x^{\uparrow}$ denotes the non-decreasing reordering of $x$. We write $e \in \mathbb{R}^n$ to denote $e = (1, \ldots, 1)$.*

*Given $x, y \in \mathbb{R}^n$ we say that $x$ is **submajorized** by $y$, denoted $x \prec_w y$, if*

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}, \qquad k = 1, \ldots, n,$$

*and $x$ is **majorized** by $y$, denoted $x \prec y$, if $x \prec_w y$ and $\mathrm{Tr}(x) = \mathrm{Tr}(y)$, where*

$$\mathrm{Tr}(x) = \sum_{j=1}^{n} x_j.$$

*A matrix $A \in M_n(\mathbb{R})$ is **row stochastic** if $A_{kj} \geq 0$ for all $k, j$ and*

$$\sum_{j=1}^{n} A_{kj} = 1, \qquad k = 1, \ldots, n.$$

*The matrix $A$ is **doubly stochastic** if both $A$ and $A^\top$ are row stochastic. Equivalently, $A$ is doubly stochastic if $A_{kj} \geq 0$ for all $k, j$ and*

$$\sum_{j=1}^{n} A_{kj} = 1, \quad k = 1, \ldots, n \qquad \text{and} \qquad \sum_{k=1}^{n} A_{kj} = 1, \quad j = 1, \ldots, n.$$

*We write*

$$\mathrm{DS}(n) = \{A \in M_n(\mathbb{R}) : \ A \text{ is doubly stochastic}\}.$$

If $A, B \in M_n(\mathbb{R})$ are doubly stochastic, then so is $AB$ (Exercise B.1.8). A basic observation about majorization is that it does not depend on the actual order of the entries of $x$ and $y$, as it is defined in terms of reorderings of said entries.

Majorization in a sense measures how "concentrated" the entries of $x$ are. Indeed, when the entries of $x$ are non-negative we always have

$$(\mathrm{Tr}(x))\, e \prec x \prec (\mathrm{Tr}(x), 0, \ldots, 0)$$

(Exercise B.1.1).

**B.1.2. Definition.**

*Given $\sigma \in \mathbb{S}_n$, we associate with it a permutation matrix $P_\sigma \in M_n(\mathbb{R})$ by considering it as the operator*

$$P_\sigma x = (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

*Explicitly,*

$$(P_\sigma)_{kj} = \begin{cases} 1, & k = \sigma(j) \\ 0, & \text{otherwise} \end{cases}$$

*A linear map $T : \mathbb{R}^n \to \mathbb{R}^n$ is a **T-Transform** if $T$ is of the form $Tx = tx + (1-t)P_\sigma x$, where $t \in [0, 1]$ and $\sigma = (k\ j)$ is a transposition.*

When $T$ is a $T$-transform, the vector $Tx$ is obtained from $x$ by replacing the $j^{\text{th}}$ entry of $x$ with $tx_j + (1-t)x_k$, and the $k^{\text{th}}$ entry with $tx_k + (1-t)x_j$. For example, when $n = 3$ and $\sigma = (1\ 3)$,

$$T = \begin{bmatrix} t & 0 & 1-t \\ 0 & 1 & 0 \\ 1-t & 0 & t \end{bmatrix}$$

It is apparent that $T \in \mathrm{DS}(3)$, and this is true in general (Exercise B.1.7).

Majorization is almost a partial order. It is not really antisymmetric (as it doesn't see permutations), but we do have that

$$x \prec y \text{ and } y \prec x \iff y = P_\sigma x \text{ for some permutation } \sigma \tag{B.1}$$

(Exercise B.1.3).

*B.1.3. Remark*. Since convex combinations will appear often in what is coming, we will use the following shortcut. A convex combination of $x_1, \dots, x_m$ is $\sum_j t_j x_j$, where $t_j \geq 0$ for all $j$ and $\sum_j t_j = 1$. These two conditions can be summarized by saying that $t \prec e_1$, where $t = (t_1, \dots, t_m)$ (Exercise B.1.5).

There are many equivalent characterizations of majorization. Here we list a few.

***B.1.4. Proposition.***

*Let $x, y \in \mathbb{R}^n$. The following statements are equivalent:*

*(i)* $x \prec y$;

*(ii)* $\displaystyle\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow$ *and* $\displaystyle\sum_{j=1}^k x_j^\uparrow \geq \sum_{j=1}^k y_j^\uparrow$, $k = 1, \dots, n$;

*(iii) for any* $f : [a,b] \to \mathbb{R}$ *convex, where* $x_j, y_j \in [a,b]$ *for all* $j$, $\operatorname{Tr} f(x) \leq \operatorname{Tr} f(y)$ ;

*(iv) for all* $t \in \mathbb{R}$, $\displaystyle\sum_{j=1}^n |x_j - t| \leq \sum_{j=1}^n |y_j - t|$;

*(v) there exists* $A \in \mathrm{DS}(n)$ *with* $x = Ay$;

*(vi)* $x \in \operatorname{conv}\{P_\sigma y :\ \sigma \in \mathbb{S}_n\}$;

*(vii) there exist $T$-transforms* $T_1, \dots, T_r$ *such that* $x = T_r \cdots T_1 y$.

*Proof.* The proof that (i) and (ii) are equivalent will be left as an exercise (Exercise B.1.10).

We will prove (vi) $\implies$ (iii), (iii) $\implies$ (iv), and then the cycle

$$\text{(i)} \implies \text{(vii)} \implies \text{(vi)} \implies \text{(v)} \implies \text{(iv)} \implies \text{(i)}.$$

(vi) $\implies$ (iii) We have $x = \sum_r \alpha_r P_{\sigma_r}$ for $\alpha \prec e_1$ and permutations $\{\sigma_r\}$. Then

$$\begin{aligned}\operatorname{Tr} f(x) &= \sum_{j=1}^{n} f\Big(\Big[\sum_r \alpha_r P_{\sigma_r} y\Big]_j\Big) = \sum_{j=1}^{n} f\Big(\sum_r \alpha_r [P_{\sigma_r} y]_j\Big) \\ &\leq \sum_{j=1}^{n}\sum_r \alpha_r f([P_{\sigma_r} y]_j) = \sum_{r=1}^{n} \alpha_r \sum_j f([P_{\sigma_r} y]_j) \\ &= \sum_r \operatorname{Tr}(f(y)) = \operatorname{Tr}(f(y)).\end{aligned}$$

(iii) $\implies$ (iv) The function $s \longmapsto |s-t|$ is convex.

(i) $\implies$ (vii) When $\sigma$ is a transposition, $P_\sigma$ is a $T$-transform, since $P_\sigma = tI + (1-t)P_\sigma$ for $t = 0$. This means that permutations are products of $T$-transforms and hence we do not lose generality if we permute $x, y$ or both; which means that we can assume that $x = x^\downarrow$, $y = y^\downarrow$.

We will proceed by induction on $n$. We have $y_n \leq x_1 \leq y_1$. This guarantees that $k = \min\{j : y_j \leq x_1\}$ exists. Choose $t \in [0,1]$ with $x_1 = ty_1 + (1-t)y_k$, and put $T_1 = tI + (1-t)P_{(1\ k)}$. So

$$T_1 y = (x_1, y_2, \ldots, y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, \ldots, y_n).$$

We now want to show that $x' \prec y'$, where

$$x' = (x_2, \ldots, x_n), \qquad y' = (y_2, \ldots, y_{k-1}, (1-t)y_1 + ty_k, y_{k+1}, \ldots, y_n).$$

We know from the choice of $k$ that $x_1 < y_{k-1}$. Then, when $\ell \leq k-1$,

$$\sum_{j=2}^{\ell} x_j \leq (\ell - 1) y_{k-1} \leq \sum_{j=2}^{\ell} y_j.$$

And when $k \leq \ell \leq n$,

$$\begin{aligned}\sum_{j=2}^{\ell} x_j &= \sum_{j=1}^{\ell} x_j - x_1 \\ &\leq \sum_{j=1}^{\ell} y_j - x_1 = \sum_{j=1}^{\ell} y_j - ty_1 - (1-t)y_k \\ &= \sum_{j=2}^{k-1} y_j + (1-t)y_1 + ty_k + \sum_{j=k+1}^{\ell} y_j.\end{aligned}$$

When $\ell = n$ the inequality above is an equality, so we have shown that $x' \prec y'$. By inductive hypothesis, there exist $T$-transform $T_2, \ldots, T_r$ with

$x' = T_r \cdots T_2 y'$. As $x$ and $T_1 y$ have the same first entry, we can consider $T_2, \ldots, T_r$ as acting on $\mathbb{R}^n$ if we force them to fix the first entry. Hence we get $x = T_r \cdots T_1 y$.

(vii) $\implies$ (vi) We know that $x = T_r \cdots T_1 y$. Write $T_j = tI + (1-t)P_{\eta_j}$. Then

$$T_j\Big(\sum_{r=1}^{m} \alpha_j P_{\sigma_r} y\Big) = \sum_{r=1}^{m} t\alpha_j P_{\sigma_r} y + \sum_{r=1}^{m} (1-t)\alpha_j P_{\eta_j \sigma_r} y. \tag{B.2}$$

If $\alpha \prec e_1$ then $(t\alpha, (1-t)\alpha) \prec e_1$ and so (B.2) is in $\operatorname{conv}\{P_\sigma y : \sigma \in \mathbb{S}_n\}$. Now we apply this repeatedly until we get $x \in \operatorname{conv}\{P_\sigma y : \sigma \in \mathbb{S}_n\}$.

(vi) $\implies$ (v) By hypothesis $x = \sum_r \alpha_r P_{\sigma_r} y$ for convex coefficients $\alpha_r$ and permutations $P_{\sigma_r}$. And $\sum_r \alpha_r P_{\sigma_r} \in \mathrm{DS}(n)$, as permutations are doubly stochastic and convex combinations of doubly stochastic are again doubly stochastic (Exercise B.1.6).

(v) $\implies$ (iv) Fix $t \in \mathbb{R}$. We have

$$\begin{aligned}\sum_{k=1}^{n} |x_k - t| = \sum_{k=1}^{n} |(Ay)_k - t| &= \sum_{k=1}^{n} \Big|\sum_{j=1}^{n} A_{kj} y_j - t\Big| = \sum_{k=1}^{n} \Big|\sum_{j=1}^{n} A_{kj}(y_j - t)\Big| \\ &\leq \sum_{k=1}^{n}\sum_{j=1}^{n} A_{kj}|y_j - t| = \sum_{j=1}^{n} |y_j - t|.\end{aligned}$$

(iv) $\implies$ (i) Choosing $t < \min\{x_n, y_n\}$ we get

$$\sum_{j=1}^{n} x_j - t = \sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t| = \sum_{j=1}^{n} y_j - t.$$

Thus $\displaystyle\sum_{j=1}^{n} x_j \leq \sum_{j=1}^{n} y_j$. Similarly, if $t > \max\{x_1, y_1\}$, then

$$\sum_{j=1}^{n} t - x_j = \sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t| = \sum_{j=1}^{n} t - y_j$$

and combining with the above we get $\displaystyle\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j$. Now suppose that $\sum_{j=1}^{k} x_j > \sum_{j=1}^{k} y_j$ for some $k$. Choose $t = y_k$. Then

$$\begin{aligned}\sum_{j=1}^{n} |y_j - t| &= (n-2k)t - \operatorname{Tr}(y) + 2\sum_{j=1}^{k} y_j \\ &< (n-2k)t - \operatorname{Tr}(x) + 2\sum_{j=1}^{k} x_j\end{aligned}$$

$$= \sum_{j=1}^{k} x_j - t + \sum_{j=k+1}^{n} t - x_j$$

$$\leq \sum_{j=1}^{n} |x_k - t|,$$

a contradiction. It follows that $x \prec_w y$ and, together with the equality of traces, $x \prec y$. □

**B.1.5. Definition.**

*Given $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, we denote by $(x, y) \in \mathbb{R}^{n+m}$ the vector*

$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m).$$

***B.1.6. Corollary.***

*Let $x, y \in \mathbb{R}^n$ and $w, z \in \mathbb{R}^m$ with $x \prec y$ and $w \prec z$. Then $(x, w) \prec (y, z)$ in $\mathbb{R}^{n+m}$.*

*Proof.* We know from Proposition B.1.4 that there exist doubly stochastic matrices $A \in M_n(\mathbb{R})$ and $B \in M_m(\mathbb{R})$ with $x = Ay$, $w = Bz$. Then

$$(x, y)^\top = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} (x, z)^\top$$

and the matrix is doubly stochastic. By Proposition B.1.4, $(x, y) \prec (w, z)$. □

***B.1.7. Proposition.***

*Let $x, y \in \mathbb{R}^n$. The following statements are equivalent:*

*(i)* $x \prec_w y$;

*(ii)* *there exists* $z \in \mathbb{R}^n$ *with* $x \leq z \prec y$;

*(iii)* *there exists* $w \in \mathbb{R}^n$ *with* $x \prec w \leq y$.

*Proof.* (i) $\iff$ (ii) The existence of $z$ immediately guarantees that $x \prec_w y$. So let us assume that $x \prec_w y$. We can assume that $\mathrm{Tr}(x) < \mathrm{Tr}(y)$, for otherwise we have $x \prec y$ and we can take $z = x$. Also, we will initially assume that $x = x^\downarrow$ and $y = y^\downarrow$ and we will proceed by induction. When $n = 1$ the result is trivial. For $n > 1$ we can consider two cases:

- there exists $k < n$ with $\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} y_j$. If we consider $a, b \in \mathbb{R}^k$ given by $a = (x_1, \ldots, x_k)$ and $b = (y_1, \ldots, y_k)$, we have $a \prec b$. Let $c, d \in \mathbb{R}^{n-k}$ be given by $c = (x_{k+1}, \ldots, x_n)$, $d = (y_{k+1}, \ldots, y_n)$. Then $c \prec_w d$, because the entries of $x$ and $y$ are ordered and then
$$\sum_{j=1}^{p} d_j - c_j = \sum_{j=k+1}^{k+p} y_j - x_j = \sum_{j=1}^{k+p} y_j - x_j \geq 0.$$
By inductive hypothesis there exists $z' \in \mathbb{R}^{n-k}$ with $c \leq z' \prec d$. By Corollary B.1.6, $(a, z') \prec (b, d)$. Then
$$x = (a, c) \leq (a, z') \prec (b, d) = y.$$
- $d = \min\left\{ \sum_{j=1}^{k} y_j - x_j : \; k = 1, \ldots, n \right\} > 0$. Choose $k_0$ such that $d = \sum_{j=1}^{k_0} y_j - x_j$. Let $v = x + de_1$. As $x$ is ordered non-decreasingly, so is $v$. We have $x \leq v$, and the choice of $d$ guarantees that $v \prec_w y$. Now $v$ and $y$ fall in the first case, so there exists $u \in \mathbb{R}^n$ with $v \leq u \prec y$. Then $x \leq v \leq u \prec y$.

When $x$ and $y$ are not ordered, there exist permutations $P, Q$ such that $Px = x^{\downarrow}$ and $Qy = y^{\downarrow}$. By the above there exists $z$ with $Px \leq z \prec Qy$. Then $z \prec y$ also, and $x \leq P^{-1}z \prec y$.

(i) $\iff$ (iii) Assume first that $x, y$ have their entries in non-increasing order. If $x \prec_w y$ we have
$$\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j, \quad k = 1, \ldots, n.$$
Let $v' = (y_1, \ldots, y_{n-1}, -\sum_{j=1}^{n-1} y_j + \sum_{j=1}^{n} x_j)$. Then, for $k = 1, \ldots, n-1$,
$$\sum_{j=1}^{k} v'_j \geq \sum_{j=1}^{k} y_j \geq \sum_{j=1}^{k} x_j,$$
so $x \prec_w v'$. By construction, $\sum_{j=1}^{n} v'_j = \sum_{j=1}^{n} x_j$, so $x \prec v'$. Now, for the case of arbitrarily ordered entries, the above argument shows that $x \prec v' \leq Py$ for some permutation $P$. Then, with $v = P^{-1}v'$, we have $x \prec v \leq y$.

If $x \prec v \leq y$, then let $P$ be the permutation with $Pv = v^{\downarrow}$. Then, for each $k = 1, \ldots, n$,
$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} v_j^{\downarrow} \leq \sum_{j=1}^{k} (Py)_j \leq \sum_{j=1}^{k} y_j^{\downarrow}.$$

Thus $x \prec_w y$. $\square$

### B.1.1. Exercises.

(B.1.1) Let $x \in \mathbb{R}^n$ with $x_j \geq 0$ for all $j$. Show that

$$\frac{\operatorname{Tr}(x)}{n} e \prec x \prec (\operatorname{Tr}(x), 0, \ldots, 0)$$

(B.1.2) Show that $t_1, \ldots, t_n$ are convex coefficients (that, is $t_j \geq 0$ for all $j$ and $\sum_j t_j = 1$) if and only if $t \prec e_1$, where $t = (t_1, \ldots, t_n)$.

(B.1.3) Prove (B.1), that is $x \prec_w y$ and $y \prec_w x$ if and only if there exists a permutation $\sigma$ with $y = P_\sigma x$.

(B.1.4) Let $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Show that if $x \prec y$ then $\alpha x \prec \alpha y$.

(B.1.5) Show that $A \in \mathrm{DS}(n)$ if and only if $A$ has non-negative entries, $Ae = e$, and $A^\top e = e$.

(B.1.6) Show that $\mathrm{DS}(n)$ is convex.

(B.1.7) Show that a $T$-transform is doubly stochastic.

(B.1.8) Show that if $A, B \in \mathrm{DS}(n)$ then $AB \in \mathrm{DS}(n)$.

(B.1.9) Let $A \in M_n(\mathbb{R})$ be doubly stochastic.

  (i) Show that $A$ has at least $n$ nonzero entries.

  (ii) Show that $A$ has precisely $n$ nonzero entries if and only if $A$ is a permutation.

(B.1.10) Show that (ii) and (i) in Proposition B.1.4 are equivalent.

(B.1.11) Show directly that (i) $\Longrightarrow$ (iv) in Proposition B.1.4

(B.1.12) Let $x, y \in \mathbb{R}^n$. Show that the following statements are equivalent:

  (i) $x \prec_w y$;

  (ii) $\operatorname{Tr} f(x) \leq \operatorname{Tr} f(y)$ for all $f$ convex and non-decreasing;

  (iii)

$$\sum_{j=1}^n (x_j - t)^+ \leq \sum_{j=1}^n (y_j - t)^+, \qquad t \in \mathbb{R}. \tag{B.3}$$

## B.2. Some Combinatorics

We will need a tiny bit of combinatorics, concretely the **Matching Problem**. We have two sets $D = \{d_1, \ldots, d_n\}$ and $H = \{h_1, \ldots, h_n\}$, and a relation $R \subset D \times H$. The problem asks whether there exists a bijection $f : D \to H$

such that its graph is contained in $R$. This is called the "matching problem" (historically, the sets were "boys" and "girls"); we think of $D$ and $H$ as sets of doctors and hospitals, respectively, and $(d_k, h_j) \in R$ means that doctor $d_k$ lives in the area of hospital $h_j$. The desired bijection of the problem gives a way to match each doctor with a unique hospital in their area. Given $k$, let $H_k = \{h_j : (d_k, h_j) \in R\}$ (that is, $H_k$ is the set of hospitals in the area where doctor $d_k$ is). Given $k_1, \ldots, k_r$, we denote $H_{k_1,\ldots,k_r} = \bigcup_{s=1}^r H_{k_s}$; this is the set of hospitals in the area of at least one of the doctors $d_{k_1}, \ldots, d_{k_r}$. An obvious necessary condition for the problem to have a solution is that $|H_{k_1,\ldots,k_r}| \geq r$ for all choices $k_1, \ldots, k_r$ (we can only match $r$ doctors to distinct hospitals if they collectively are in the area of at least $r$ hospitals). The surprising thing is that this necessary condition is also sufficient:

***B.2.1. Theorem. (Hall)***

*With the notation above, a compatible matching between $D$ and $H$ is possible if and only if*

$$|H_{k_1,\ldots,k_r}| \geq r \qquad \text{for all choices } k_1, \ldots, k_r \tag{B.4}$$

*Proof.* It was mentioned above that the condition is necessary. So it remains to assume (B.4) and show that the bijection exists. We proceed by induction on $n$. When $n = 1$ we have $|H_1| \geq 1$, which guarantees that the doctor is in the only hospital's area. For the inductive step, we will subdivide the conditions into two possibilities. Suppose first that

$$|H_{k_1,\ldots,k_r}| \geq r + 1 \qquad \text{for all choices } k_1, \ldots, k_r,\ 1 \leq r < n.$$

Match $d_1$ with any hospital in $H_1$; after removing this hospital from all the sets $H_{k_1,\ldots,k_r}$ their cardinality decreases at most by one, so (B.4) still holds, and the inductive hypothesis matches the remaining $n-1$ doctors with the remaining $n-1$ hospitals.

The second possibility is that

$$\text{there exist indices } k_1, \ldots, k_r, \text{ with } r < n, \text{ such that } |H_{k_1,\ldots,k_r}| = r. \tag{B.5}$$

This means that the doctors $d_{k_1}, \ldots, d_{k_r}$ have exactly $r$ hospitals in their areas. By the inductive hypothesis, these doctors and hospitals can be matched. We are left with $n-r$ doctors and $n-r$ hospitals, and a priori we don't know who is in who's area. But if $s$ of these doctors covered the areas of less than $s$ of the remaining hospitals, together with the first $r$ doctors we would have $r + s$ doctors that cover the area of less than $r + s$ hospitals (the condition (B.5) guarantees that the first $k$ doctors are not in the area of any of the remaining $n - r$ hospitals), contradicting (B.4). So (B.4) is satisfied for the remaining doctors and hospitals, who then can be matched by inductive hypothesis. $\square$

**B.2.2. Definition.**

*Let $A \in M_n(\mathbb{R})$. A* ***diagonal*** *of $A$ is $\{A_{1,\sigma(1)}, \ldots, A_{n,\sigma(n)}\}$ for some $\sigma \in \mathbb{S}_n$. A matrix $B \in M_{p,q}(\mathbb{R})$ is a* ***submatrix*** *of $A$ if there exist $\{k_1, \ldots, k_p\}$ and $\{j_1, \ldots, j_q\} \subset \{1, \ldots, n\}$ such that*

$$B_{r,s} = A_{k_r,j_s}, \qquad r = 1, \ldots, p, \ s = 1, \ldots, q.$$

*We use the notation $A[k_1, \ldots, k_p; j_1, \ldots, j_q]$ to represent the submatrix $B$ above. We also use the notation $A\{k_1, \ldots, k_p; j_1, \ldots, j_q\}$ to denote the submatrix of $A$ obtained after removing said rows and columns.*

A "diagonal" for $A$, as defined above, amounts to choosing a single element from each row, without repeating columns.

***B.2.3. Corollary. (König–Frobenius)***

*Let $A \in M_n(\mathbb{R})$. The following statements are equivalent:*

*(i) every diagonal of $A$ has a zero element;*

*(ii) there exist $p, q \in \mathbb{N}$ with $p + q > n$ and such that $A$ has a $p \times q$ zero submatrix.*

*Proof.* Exercise B.2.1. □

***B.2.4. Corollary.***

*Let $A \in M_n(\mathbb{R})$ be doubly stochastic. Then $A$ has a nonzero diagonal.*

*Proof.* Suppose that every diagonal of $A$ has a zero. By Corollary B.2.3 there exist $K, J \subset \{1, \ldots, n\}$ with $|K| + |J| > n$ and $A_{kj} = 0$ for all $k \in K$, $j \in J$. By pushing these rows and columns to the top and left respectively, we obtain permutations $P, Q$ such that

$$PAQ = \begin{bmatrix} 0_{|K|,|J|} & B \\ C & D. \end{bmatrix}$$

As $A$ is doubly stochastic, so is $PAQ$. Then the columns of $C$ add to 1, and the rows of $B$ add to 1; also, $D$ is non-negative. As $C$ has $|J|$ columns and $B$ has $|K|$ rows, this means that the entries of $A$ add to at least $|K| + |J| > n$. This is a contradiction, for the entries of an $n \times n$ doubly stochastic matrix add to $n$ (as it has $n$ rows, each adding to 1). It follows that $A$ has a nonzero diagonal. □

### B.2.1. Exercises.

(B.2.1) Prove Corollary B.2.3.

## B.3. Birkhoff's Theorem and Convex Functions

Birkhoff's Theorem for doubly stochastic matrices [Bir46], stated and proven below, has the weird connection (for this book's author, at least) that it was published in an obscure Argentinean journal. For all we know, in the 1940s several American mathematicians, including Marshall Stone, Antoni Zygmund, and George Birkhoff (Garrett Birkhoff's father) visited Argentina to build ties with the mathematical community there, and to recruit talent. It is conceivable that Birkhoff Jr. sent the paper to the journal of a small department of mathematics in Argentina to contribute to the cause of expanding mathematical research. We do not know if he actually visited Argentina. Hopefully some reader will confirm/correct the above account.

***B.3.1. Theorem. (Birkhoff)***

*The set* $\mathrm{DS}(n)$ *is convex, every double stochastic matrix is a convex combination of permutations, and the extreme point of* $\mathrm{DS}(n)$ *are the permutations.*

*Proof.* That convex combinations of doubly stochastic matrices are doubly stochastic is straightforward (Exercise B.1.6). It remains to show that the extreme points are the permutations and that their convex hull is all of $\mathrm{DS}(n)$. Let us first show that permutations are extreme. For this, suppose $P_\sigma = tA + (1-t)B$ with $A, B$ doubly stochastic and $t \in [0,1]$. The entries of $P_\sigma$ can only be 0 or 1. When $(P_\sigma)_{kj} = 0$, as the entries of $A$ and $B$ are non-negative we get that $A_{kj} = B_{kj} = 0$. And when $(P_\sigma)_{kj} = 1$, we have

$$1 = tA_{kj} + (1-t)B_{kj}.$$

As $0 \leq A_{kj}, B_{kj} \leq 1$, this forces $A_{kj} = B_{kj} = 1$. Thus $A = B = P_\sigma$ and $P_\sigma$ is extreme.

Now we show that every doubly stochastic matrix is a convex combination of permutations. We will proceed by induction on $r$, the number of positive entries of $A$; we necessarily have $r \geq n$. When $r = n$, as $A$ has a nonzero element in each row and column, the only possibility is that each row and column has an entry equal to 1 and the rest equal to 0; and then $A$ is a permutation. Now suppose that $A$ has $r+1$ positive entries. By

Corollary B.2.4 there exists $\sigma \in \mathbb{S}_n$ such that $\{A_{1,\sigma(1)}, \ldots, A_{n,\sigma(n)}\}$ are all positive. Let $a = \min\{A_{1,\sigma(1)}, \ldots, A_{n,\sigma(n)}\}$. If $a = 1$, then the whole diagonal would consist of 1 and then $A$ would be a permutation. Otherwise, $a < 1$ and $r + 1 > n$. Let

$$B = \frac{A - aP_\sigma}{1 - a}.$$

Then $B$ is doubly stochastic, for

$$B = \frac{1}{1-a}A + \frac{a}{1-a}P_\sigma.$$

Let $h$ such that $a = A_{h,\sigma(h)}$. Then $B_{h,\sigma(h)} = \frac{a-a}{1-a} = 0$. When $A_{kj} = 0$ we know that $j \neq \sigma(k)$, so $(P_\sigma)_{kj} = 0$; thus $B_{kj} = 0$. It follows that $B$ has at most $r$ positive entries. By inductive hypothesis there exists $t \in \mathbb{R}^s$ with $t \prec e_1$ and $P_{\sigma_1}, \ldots, P_{\sigma_s}$ such that $B = \sum_j t_j P_{\sigma_j}$. Then

$$A = aP_\sigma + (1-a)B = aP_\sigma + \sum_{j=1}^{s}(1-a)t_j P_{\sigma_j} \in \operatorname{conv}\{P_\eta : \ \eta \in \mathbb{S}_n\}. \quad \square$$

*B.3.2. Remark*. The proof of Birkhoff's Theorem above is more or less the canonical one, and its beauty, together with the background it requires (Hall's Theorem) should be appreciated. With the machinery we have available, though, a fairly direct proof is possible. Indeed, as in the first paragraph of Theorem B.3.1 we have that $\mathrm{DS}(n)$ is convex and that the permutations are extreme; suppose on top of this that we show that any extreme point is a permutation (we will do this in the next paragraph). The properties defining doubly stochasticity are continuous, and so $\mathrm{DS}(n)$ is closed; being a closed bounded subset of a finite-dimensional normed space, it is compact. By Carathéodory (Theorem 7.5.12)

$$\mathrm{DS}(n) = \operatorname{conv}\operatorname{Ext}\mathrm{DS}(n) = \operatorname{conv}\{P_\sigma : \ \sigma \in \mathbb{S}_n\}.$$

So it remains to show that $\operatorname{Ext}\mathrm{DS}(n) = \{P_\sigma : \ \sigma \in \mathbb{S}_n\}$. If $A \in \mathrm{DS}(n)$ has an entry with $A_{rs} \in (0,1)$ for some $r, s$, then $A$ is not extreme by Exercise B.3.1 (properly, the exercise considers the case of $\mathrm{DSS}(n)$, but the same argument works for $\mathrm{DS}(n)$). So every extreme point is a permutation (as it has a single 1 in every row and column), and conversely every permutation is extreme.

**B.3.3. Definition.**

*If $\varphi : \mathbb{R} \to \mathbb{R}$ and $x \in \mathbb{R}^n$, we denote by $\varphi(x) \in \mathbb{R}^n$ the element $\varphi(x) = (\varphi(x_1), \ldots, \varphi(x_n))$.*

*For $x, y \in \mathbb{R}^n$ we take $x \leq y$ to mean $x_j \leq y_j$ for all $j$.*

**B.3.4. Proposition.**

*Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex, $x, y \in \mathbb{R}^n$.*

(i) *If $x \prec y$, then $\varphi(x) \prec_w \varphi(y)$.*

(ii) *If moreover $\varphi$ is monotone increasing and $x \prec_w y$, then $\varphi(x) \prec_w \varphi(y)$.*

*Proof.*

(i) Suppose that $x \prec y$. For any permutation matrix $P$ we have $\varphi(Px) = P\varphi(x)$. By Proposition B.1.4 there exists $A \in \mathrm{DS}(n)$ with $x = Ay$. By Theorem B.3.1 there exist $t \in \mathbb{R}^r$ with $t \prec e_1$ and permutation matrices $P_1, \ldots, P_r$ such that $A = \sum_j t_j P_j$. Then

$$\varphi(x) \le \sum_{j=1}^{r} t_j \varphi(P_j y) = \sum_{j=1}^{r} t_j P_j \varphi(y).$$

By Proposition B.1.4, $\sum_{j=1}^{r} t_j P_j \varphi(y) \prec \varphi(y)$. Then

$$\varphi(x) \le \sum_{j=1}^{r} t_j P_j \varphi(y) \prec \varphi(y),$$

which gives $\varphi(x) \prec_w \varphi(y)$.

(ii) Now suppose that $x \prec_w y$ and that $\varphi$ is also monotone increasing. By Proposition B.1.7 there exists $z \in \mathbb{R}^n$ with $x \le z \prec y$. Using that $\varphi$ is monotone increasing for the first inequality and the first part of the proof for the submajorization,

$$\varphi(x) \le \varphi(z) \prec_w \varphi(y).$$

This gives $\varphi(x) \prec_w \varphi(y)$.

□

The proof below looks simple enough, but we encourage the reader to trace back the results used to prove it.

**B.3.5. Corollary.**

*Let $a_1, \ldots, a_n, b_1, \ldots, b_n \subset (0, \infty)$ such that $\prod_{j=1}^{n} a_j \le \prod_{j=1}^{n} b_j$. Then $\sum_{j=1}^{n} a_j \le \sum_{j=1}^{n} b_j$.*

*Proof.* Applying log we have

$$\sum_{j=1}^{n} \log a_j \leq \sum_{j=1}^{n} \log b_j.$$

Now Proposition B.3.4, with $\varphi(t) = e^t$, gives us

$$\sum_{j=1}^{n} a_j \leq \sum_{j=1}^{n} b_j. \qquad \square$$

### B.3.1. Exercises.

(B.3.1) A matrix $A \in M_n(\mathbb{C})$ is **doubly substochastic** if $A_{kj} \geq 0$ for all $k, j$ and

$$\sum_{j=1}^{n} A_{kj} \leq 1, \qquad k = 1, \ldots, n$$

and

$$\sum_{k=1}^{n} A_{kj} \leq 1, \qquad j = 1, \ldots, n.$$

Let $A \in \mathrm{DSS}(n)$ such that there exist $k, j$ with $A_{kj} \in (0, 1)$. Show that $A$ is not extreme.

(B.3.2) Show that the set $\mathrm{DSS}(n)$ of all doubly substochastic $n \times n$ matrices is convex, and its extreme points are those $A \in M_n(\mathbb{C})$ with at most an entry 1 in each row, and all other entries equal to zero (in particular, the zero matrix is an extreme point of $\mathrm{DSS}(n)$).

(B.3.3) Show that $B$ is doubly substochastic if and only if there exists $A \in \mathrm{DS}(n)$ with $0 \leq B_{kj} \leq A_{kj}$ for all $k, j$.

(B.3.4) Let $x, y \in \mathbb{C}^n$ with non-negative coordinates. Show that $x \prec_w y$ if and only if $x = By$ for some $B \in \mathrm{DSS}(n)$.

APPENDIX C

# Lidskii's Theorem

The goal here is to prove Lidskii's formula (10.26): for any $T \in \mathcal{T}(\mathcal{H})$, if $\{\lambda_k\}$ denote the eigenvalues of $T$, counting multiplicities, then

$$\operatorname{Tr}(T) = \sum_k \lambda_k. \tag{C.1}$$

For all the simplicity in the statement, there is no elementary proof in infinite dimension. The proof in [Sim05], that we mostly follow, depends on results that were not included in that text. Here we have strived to write a self-contained proof. We do recommend [Sim05] for a longer discussion of the origins and alternative proofs of the theorem.

## C.1. Antisymmetric Tensor Products and the Determinant

We begin by extending the definition of the tensor product of Hilbert spaces (Section 13.1) to products of $n$ spaces instead of two.

**C.1.1. Definition.**

*Let $\mathcal{H}$ be a Hilbert space, and $n \in \mathbb{N}$. The $n$-fold tensor product of $\mathcal{H}$, denoted by $\bigotimes^n \mathcal{H}$, is the completion of the span of the linear maps $\xi_1 \otimes \cdots \otimes \xi_n : ML(\mathcal{H}^n) \to \mathbb{C}$, where for a multilinear map $\phi : \mathcal{H}^n \to \mathbb{C}$*

*we have*

$$(\xi_1 \otimes \cdots \otimes \xi_n)(\phi) = \phi(\xi_1, \ldots, \xi_n),$$

*with the inner product induced by*

$$\langle \xi_1 \otimes \cdots \otimes \xi_n, \eta_1 \otimes \cdots \otimes \eta_n \rangle = \langle \xi_1, \eta_1 \rangle \cdots \langle \xi_n, \eta_n \rangle.$$

*Given $\xi_1, \ldots, \xi_n \in \mathcal{H}$, their* ***antisymmetric product*** *(also:* ***exterior product****) is $\xi_1 \wedge \cdots \wedge \xi_n$ given by*

$$\xi_1 \wedge \cdots \wedge \xi_n = \frac{1}{\sqrt{n!}} \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn} \sigma \, \xi_{\sigma(1)} \otimes \cdots \otimes \xi_{\sigma(n)}.$$

*We write*

$$\bigwedge^n \mathcal{H} = \overline{\operatorname{span}}^{\|\cdot\|} \{\xi_1 \wedge \cdots \wedge \xi_n : \ \xi_1, \ldots, \xi_n \in \mathcal{H}\}.$$

*Given $T \in \mathcal{B}(\mathcal{H})$, we denote by $\bigwedge^n(T) \in \mathcal{B}(\bigwedge^n(\mathcal{H}))$ the linear operator induced by*

$$\bigwedge^n(T)(\xi_1 \wedge \cdots \wedge \xi_n) = T\xi_1 \wedge \cdots \wedge T\xi_n.$$

When $T \in \mathcal{B}(\mathbb{C}^n) = M_n(\mathbb{C})$, we can triangularize it via an orthonormal basis (Proposition 1.7.14). So there exists an orthonormal basis $\{\xi_1, \ldots, \xi_n\}$ of $\mathbb{C}^n$ such that $\langle T\xi_k, \xi_k \rangle = \langle \lambda_k(T)\xi_k, \xi_k \rangle = \lambda_k(T)$. Then

$$\operatorname{Tr}(T) = \sum_{k=1}^n \langle T\xi_k, \xi_k \rangle = \sum_{k=1}^n \langle \lambda_k(T)\xi_k, \xi_k \rangle = \sum_{k=1}^n \lambda_k(T).$$

Also,

$$\det(I+T) = \sum_{\sigma \in \mathbb{S}_n} (\operatorname{sgn} \sigma) \, (I+T)_{\sigma(1),1} \cdots (I+T)_{\sigma(n),n}. \tag{C.2}$$

The determinant does not depend on the basis we use to represent the operator (because $\det(VTV^{-1}) = \det T$), so we might as well use a basis that triangularizes $I+T$. Then $(I+T)_{kj} = 0$ whenever $k > j$. Looking at (C.2), the product can be nonzero only when $\sigma(1) \leq 1$, so $\sigma(1) = 1$; we also need $\sigma(2) \leq 2$, so the only option is $\sigma(2) = 2$. Continuing this way, the only term that is nonzero is the one for $\sigma = \mathrm{id}$. Thus

$$\det(I+T) = \prod_{k=1}^n (1 + \lambda_k(T)). \tag{C.3}$$

We have that

$$\left\{ \frac{1}{\sqrt{k!}} \, \xi_{j_1} \wedge \cdots \wedge \xi_{j_k} : \ j_1 < \cdots < j_k \right\} \tag{C.4}$$

is an orthonormal basis for $\bigwedge^k \mathcal{H}$ (Exercise C.1.1). Then

$$\operatorname{Tr}(\bigwedge^k(T)) = \sum_{j_1 < \cdots < j_k} \langle T\xi_{j_1} \wedge \cdots \wedge T\xi_{j_k}, \xi_{j_1} \wedge \cdots \wedge \xi_{j_k} \rangle$$

$$= \frac{1}{k!} \sum_{j_1 < \cdots < j_k} \sum_{\sigma, \sigma' \in \mathbb{S}_k} \operatorname{sgn}\sigma \, \operatorname{sgn}\sigma' \prod_{t=1}^{k} \langle T\xi_{j_{\sigma(t)}}, \xi_{j_{\sigma'(t)}} \rangle$$

$$= \sum_{j_1 < \cdots < j_k} \sum_{\sigma} \operatorname{sgn}\sigma \prod_{t=1}^{k} \langle T\xi_{j_t}, \xi_{j_{\sigma(t)}} \rangle.$$

Because $T$ is triangular for the basis $\{\xi_k\}$, the product $\prod_{t=1}^{k} \langle T\xi_{j_t}, \xi_{j_{\sigma(t)}} \rangle$ is nonzero only when $j_t \leq j_{\sigma(t)}$ for all $t$. Thus $j_k \leq j_{\sigma(k)}$, which forces $\sigma(k) = k$; we also need $j_{k-1} \leq j_{\sigma(k-1)}$, so $\sigma(k-1) \geq k-1$ and we already have $\sigma(k) = k$, so $\sigma(k-1) = k-1$. Continuing in this fashion, the product in nonzero only when $\sigma = \mathrm{id}$. In that case, $\langle T\xi_{j_t}, \xi_{j_t} \rangle = \lambda_{j_t}(T)$. Hence

$$\operatorname{Tr}(\textstyle\bigwedge^k(T)) = \displaystyle\sum_{j_1 < \cdots < j_k} \lambda_{j_1}(T) \cdots \lambda_{j_k}(T), \tag{C.5}$$

so in particular

$$\operatorname{Tr}(\textstyle\bigwedge^n(T)) = \det T,$$

and

$$\begin{aligned} \sum_{k=0}^{n} \operatorname{Tr}(\textstyle\bigwedge^k(T)) &= \displaystyle\sum_{k=0}^{n} \sum_{j_1 < \cdots < j_k} \lambda_{j_1}(T) \cdots \lambda_{j_k}(T) \\ &= \prod_{k=1}^{n} (1 + \lambda_k(T)) = \det(I+T). \end{aligned} \tag{C.6}$$

*C.1.2. Remark.* This whole antisymmetric business looks cumbersome, and to some extent it is. But we will get mileage out of it. And, also, not everything is as bad as it looks; sometimes, the exterior product simplifies things. For instance, when $\mathcal{H} = \mathbb{C}^n$ and $T \in M_n(\mathbb{C})$, we have that $\dim \bigwedge^n \mathcal{H} = 1$ and $\bigwedge^n T$ is the operator of scalar multiplication by $\det T$ (Exercise C.1.3).

The idea going forward is to use (C.6) to define $\det(I+T)$ for $T \in \mathcal{T}(\mathcal{H})$, but we still have to see that the expression actually makes sense. That is the reason for the next two lemmas.

***C.1.3. Lemma.***

*Let $T \in \mathcal{T}(\mathcal{H})$ and $k \in \mathbb{N}$. Then $\bigwedge^k(T) \in \mathcal{T}(\bigwedge^k \mathcal{H})$ and $\| \bigwedge^k(T)\|_1 \leq \frac{1}{k!} \|T\|_1^k$.*

*Proof.* As $\bigwedge^k(T)$ acts component-wise, we have $\bigwedge^k(T)^* \bigwedge^k(T) = \bigwedge^k(T^*T)$, and the same happens with the square root; thus, $|\bigwedge^k(T)| = \bigwedge^k(|T|)$. Let us check that the equality (C.5) holds for $|T|$. Since $|T|$ is positive, by Remark 10.6.13 we can choose an orthonormal basis $\{\xi_j\}$ of $\mathcal{H}$ such that

$|T|\xi_j = \lambda_j(|T|)\xi_j$ for all $j$. Then, with the orthonormal basis from (C.4),

$$\begin{aligned}\operatorname{Tr}(\textstyle\bigwedge^k(|T|) &= \frac{1}{k!}\sum_{j_1<\cdots<j_k}\langle |T|\xi_{j_1}\wedge\cdots\wedge|T|\xi_{j_k}, \xi_{j_1}\wedge\cdots\wedge\xi_{j_k}\rangle \\ &= \frac{1}{k!}\sum_{j_1<\cdots<j_k}\lambda_{j_1}(|T|)\cdots\lambda_{j_k}(|T|)\,\|\xi_{j_1}\wedge\cdots\wedge\xi_{j_k}\|^2 \\ &= \sum_{j_1<\cdots<j_k}\lambda_{j_1}(|T|)\cdots\lambda_{j_k}(|T|).\end{aligned} \tag{C.7}$$

Now, using Exercise C.1.2,

$$\begin{aligned}\|\textstyle\bigwedge^k(T)\|_1 = \operatorname{Tr}(\bigwedge^k(|T|)) &= \sum_{j_1<\cdots<j_k}\lambda_{j_1}(|T|)\cdots\lambda_{j_k}(|T|) \\ &\le \frac{1}{k!}\sum_{j_1,\ldots,j_k=1}^{\infty}\lambda_{j_1}(|T|)\cdots\lambda_{j_k}(|T|) \\ &= \frac{1}{k!}\Big(\sum_{j=1}^{\infty}\lambda_j(|T|)\Big)^k = \frac{1}{k!}\|T\|_1^k.\end{aligned}$$

□

We can now formally define the determinant on these particular Fredholm operators:

**C.1.4. Definition.**

*Let $T \in \mathcal{T}(\mathcal{H})$. The* ***determinant*** *of $I+T$ is defined as*

$$\det(I+T) = \sum_{k=0}^{\infty}\operatorname{Tr}(\textstyle\bigwedge^k(T)).$$

Of course we need to check that the series in the definition converges.

***C.1.5. Lemma.***

*Let $S, T \in \mathcal{T}(\mathcal{H})$. The series*

$$\det(I+S+zT) = \sum_{k=0}^{\infty}\operatorname{Tr}(\textstyle\bigwedge^k(S+zT))$$

*defines an entire function that satisfies*

$$|\det(I+S+zT)| \le e^{\|S\|_1}\exp(\|T\|_1\,|z|). \tag{C.8}$$

*Also, given any $\varepsilon > 0$ there exists a constant $C_\varepsilon$ such that*

$$|\det(I+zT)| \le C_\varepsilon\exp(\varepsilon\,|z|). \tag{C.9}$$

*Proof.* The terms on the series are polynomials on $z$ (the $k^{\text{th}}$ term has degree $k$) Using Proposition 10.7.5 and Lemma C.1.3 we have

$$\begin{aligned}\sum_{k=0}^{\infty} |\operatorname{Tr}(\textstyle\bigwedge^k(S+zT))| &\le \sum_{k=0}^{\infty} \operatorname{Tr}(\textstyle\bigwedge^k(|S+zT|)) \\ &\le \sum_{k=0}^{\infty} \frac{1}{k!}\,\|S+zT\|_1^k = \exp(\|S+zT\|_1) \\ &\le e^{\|S\|_1}\,\exp(\|T\|_1\,|z|),\end{aligned}$$

which shows both that the series converges absolutely for all $z$, uniformly on bounded sets, and the inequality (C.8). By Proposition 3.7.7, the series defines a holomorphic function. Now, using (C.7)

$$|\det(I+zT)| \le \sum_{k=0}^{\infty} \sum_{j_1<\cdots<j_k} |z|\lambda_{j_1}(|T|)\cdots|z|\lambda_{j_k}(|T|) = \prod_{k=1}^{\infty}(1+|z|\,\lambda_k(|T|)).$$

Fix $\varepsilon > 0$. If we choose $k_0$ so that $\sum_{k>k_0}\lambda_k(|T|) < \frac{\varepsilon}{2}$ and we use the inequality $(1+t) \le e^t$ for $t>0$, we get

$$\begin{aligned}|\det(I+zT)| &\le \prod_{k=1}^{k_0}(1+|z|\lambda_k(|T|)) \prod_{k>k_0} e^{|z|\,\lambda_k(|T|)} \\ &\le \prod_{k=1}^{k_0}(1+|z|\lambda_k(|T|))\, e^{|z|\,\sum_{k>k_0}\lambda_k(|T|)} \le C_\varepsilon\, e^{\varepsilon\,|z|},\end{aligned}$$

by bounding the product (which is a polynomial on $|z|$) with $C_\varepsilon\, e^{\varepsilon/2\,|z|}$ for an appropriate constant $C_\varepsilon$. □

We defined the multiplicity function $\alpha_\lambda(T)$ on Remark 10.7.6.

***C.1.6. Proposition.***

*Let $S, T \in \mathcal{T}(\mathcal{H})$. Then*

*(i)* $|\det(I+S) - \det(I+T)| \le \|S-T\|_1\,\exp(\|S\|_1 + \|T\|_1 + 1)$*;*

*(ii)* $\det\big((I_T)(I+S)\big) = \det(I+S)\det(I+T)$*;*

*(iii) if $\lambda$ is a nonzero eigenvalue for $T$ with $\alpha_\lambda(T) = n$, then* $\det(I+zT)$ *has a zero of order $n$ at $z_0 = -1/\lambda$;*

*(iv) $I+T$ is invertible if and only if* $\det(I+T) \ne 0$.

*Proof.*

(i) Let $g(z) = \det(I + \frac{1}{2}(S+T) + z(S-T))$. We know that $g$ is entire by Lemma C.1.5. We have

$$\begin{aligned}|\det(I+S) - \det(I+T)| &= |g(1/2) - g(-1/2)| \\ &\le \sup\{|g'(t)| :\ t \in [-1/2, 1/2]\}.\end{aligned}$$

Fix $r > 0$. Using Cauchy's Integral Formula (3.15) on $\gamma = r\mathbb{T}$, for $|t| \le \frac{1}{2}$ we have

$$\begin{aligned}|g'(t)| &= \frac{1}{2\pi}\left|\int_\gamma \frac{g(t+w)}{w^2}\,dw\right| = \frac{1}{2\pi}\left|\int_0^{2\pi} \frac{g(t+re^{i\theta})}{r^2 e^{2i\theta}}\, ire^{i\theta}\,d\theta\right| \\ &\le \frac{1}{2\pi r}\int_0^{2\pi} \left|g(t+re^{i\theta})\right| d\theta \le \frac{1}{r}\sup\left\{|g(z)| :\ |z| \le r + \frac{1}{2}\right\}.\end{aligned}$$

Choosing $r > 0$ so that $\frac{1}{r} = \|S-T\|_1$ and using (C.8), we obtain

$$\begin{aligned}&\big|\det(I+S) - \det(I+T)\big| \\ &\quad\le \|S-T\|_1 \exp\Big(\frac{1}{2}\|S+T\|_1\Big) \exp\Big(\|S-T\|_1 \big(r + \frac{1}{2}\big)\Big) \\ &\quad= \|S-T\|_1 \exp\Big(\frac{1}{2}\|S+T\|_1 + \frac{1}{2}\|S-T\|_1 + r\|S-T\|_1\Big) \\ &\quad\le \|S-T\|_1 \exp\big(\|S\|_1 + \|T\|_1 + 1\big).\end{aligned}$$

(ii) Fix an orthonormal basis $\{\xi_k\}$ of eigenvectors for | and let $P_n$ be the orthogonal projection onto the subspace $\overline{\operatorname{span}}^{\|\cdot\|}\{\xi_1, \dots, \xi_n\}$. Then $P_n$ is finite-rank. We have

$$\begin{aligned}\|T - P_nTP_n\|_1 &\le \|T - TP_n\|_1 + \|TP_n - P_nTP_n\|_1 \\ &\le \|T - TP_n\|_1 + \|T - P_nT\|_1 \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

This is because if $T = V|T|$ is the polar decomposition of $T$ then

$$\begin{aligned}\|T - TP_n\|_1 &= \|V(|T| - |T|P_n)\|_1 \le \|\,|T| - |T|P_n\|_1 \\ &= \|(I-P_n)|T|(I-P_n)\|_1 = \sum_k \langle |T|(I-P_n)\xi_k, (I-P_n)\xi_k\rangle \\ &= \sum_{k=n+1}^{\infty} \langle |T|\xi_k, \xi_k\rangle \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

The same computation shows that $\|T - P_nT\|_1 \to 0$. As $P_nTP_n$ acts on the finite-dimensional space $P_n\mathcal{H}$, (C.6) and the discussion preceding it apply. We have, using Exercise C.1.4 and (C.6),

$$\det(I + P_nTP_n) = \sum_{k=0}^{\infty} \operatorname{Tr}(\textstyle\bigwedge^k(P_nTP_n)) = \sum_{k=0}^{n} \operatorname{Tr}(\textstyle\bigwedge^k(P_nTP_n))$$

$$= \prod_{k=1}^{n}(1+\lambda_k(P_nTP_n)) = \det_n(P_n + P_nTP_n),$$

where we use $\det_n$ to emphasize that it is the usual linear algebra determinant. Note that $P_n$ is the identity on $P_n\mathcal{H}$. By (i)

$$\begin{aligned}\det(I+S+T+ST) &= \lim_n \det(I+P_nSP_n+P_nTP_n+P_nSP_nTP_n)\\ &= \lim_n \det_n(P_n+P_nSP_n+P_nTP_n+P_nSP_nTP_n)\\ &= \lim_n \det_n((P_n+P_nSP_n)(P_n+P_nTP_n))\\ &= \lim_n \det_n(P_n+P_nSP_n)\det_n(P_n+P_nTP_n)\\ &= \det(I+S)\det(I+T).\end{aligned}$$

(iii) Since $T$ is compact, we know that its nonzero eigenvalues are isolated (Theorem 9.6.13). Fix nonzero $\lambda \in \sigma(T)$. Consider the corresponding Riesz projection $P_\lambda$, as given by Proposition 9.3.4. Then any $v \in P_\lambda\mathcal{H}$ is an eigenvector for $\lambda$ and $P_\lambda$ is finite-rank by the compactness of $T$. Since $TP_\lambda = P_\lambda T$,

$$\begin{aligned}\det(I+zTP_\lambda)\det(I+zT(I-P_\lambda)) &= \det\big(I+zTP_\lambda+zT(I-P_\lambda)\\ &\qquad + z^2TP_\lambda(I-P_\lambda)T\big)\\ &= \det(I+zTP_\lambda+zT(I-P_\lambda))\\ &= \det(I+zT).\end{aligned}$$

The operator $I+z_0T(I-P_\lambda)$ is invertible; for otherwise with $z_0 = -1/\lambda$ we would have $-1 \in \sigma(z_0T(I-P_\lambda))$, which is equivalent to $\lambda = -z_0^{-1} \in \sigma(T(I-P_\lambda)) = \sigma(T)\setminus\{\lambda\}$, a contradiction. This implies that $\det(I+z_0T(I-P_\lambda)) \neq 0$; indeed, if $S = -z_0T(I-P_\lambda)(I+z_0T(I-P_\lambda))^{-1}$, we have $(I+S)(I+z_0T(I-P_\lambda)) = I$ and then

$$1 = \det((I+S)(I+z_0T(I-P_\lambda))) = \det(I+S)\det(I+z_0T(I-P_\lambda))$$

showing that $\det(I+z_0T(I-P_\lambda)) \neq 0$. As $TP_\lambda = \lambda P_\lambda$ is finite-rank, reasoning as above with finite-rank operators we have

$$\det(I+zTP_\lambda) = \det_n(P_\lambda + z\lambda P_\lambda) = \det_n((1-zz_0^{-1})P_\lambda) = (1-zz_0^{-1})^n.$$

Thus

$$\det(I+zT) = \det(I+z_0T(I-P_\lambda))\,(1-zz_0^{-1})^n$$

has $z_0$ as a zero of order $n$.

(iv) If $I+T$ is invertible, then $\det(I+T)\neq 0$; we did this above, with the trick of taking $S = -T(I+T)^{-1}$ and using the multiplicativity of the determinant. When $I+T$ is not invertible, $-1$ is an eigenvalue for $T$ and then by (iii) we get that $\det(I+T)=0$.

□

Given $T\in\mathcal{T}(\mathcal{H})$, we proved in Lemma C.1.3 that $\bigwedge^k T$ is also trace-class. So the nonzero elements of its spectrum are eigenvalues. Here is some information about said eigenvalues.

***C.1.7. Lemma.***

*Let $T\in\mathcal{T}(\mathcal{H})$ with nonzero eigenvalues $\{\lambda_k(T):\ k\in K\}$, where $K=\{1,\ldots,m\}$ or $\mathbb{N}$. Let $k\in\mathbb{N}$. Then*

$$\prod_{r=1}^{k}\lambda_{j_r}(T),\qquad\qquad j_1<\ldots<j_k\in K \tag{C.10}$$

*are eigenvalues for $\bigwedge^k T$. When $T$ is normal, $\bigwedge^k T$ is normal and (C.10) lists all of the nonzero eigenvalues of $\bigwedge^k T$.*

*Proof.* For simplicity denote the $k$ eigenvalues $\lambda_1,\ldots,\lambda_k$. Let $V\subset\mathcal{H}$ be the (finite-dimensional) subspace of all generalized eigenvectors for $\lambda_1,\ldots,\lambda_k$ (Remark 10.7.6); this subspace is invariant for $T$. As the vectors are listed counting multiplicities and we might be missing instances (say, maybe $\lambda_{k+1}=\lambda_2$), we have that $\dim V\geq k$. But we always have that $T|_V$ has eigenvalues $\lambda_1,\ldots,\lambda_k$ (with maybe some more repetitions needed, as just mentioned). So we can apply Proposition 1.7.14 to $T|_V$, to obtain an orthonormal basis $\{\xi_j\}$ of $V$ with

$$T\xi_r=\lambda_r\xi_r+\sum_{j=1}^{r-1}\alpha_{rj}\xi_j\qquad\qquad r=1,\ldots,k \tag{C.11}$$

(this works all the way up to $r=\dim V$ which is possibly greater than $k$, but we only need to get up to $k$). Now, putting $\alpha_{rr}=\lambda_r$ for all $r$,

$$\begin{aligned}\bigwedge\nolimits^k T(\xi_1\wedge\cdots\wedge\xi_k)&=T\xi_1\wedge\cdots\wedge T\xi_k\\&=\lambda_1\xi_1\wedge\xi_{j_2}\wedge\cdots\wedge\lambda_k\xi_k\\&\qquad\qquad+\sum_{j_2=1}^{1}\sum_{j_3=1}^{2}\cdots\sum_{j_k=1}^{k-1}\left(\prod_{s=1}^{k}\alpha_{s,j_s}\right)\xi_1\wedge\cdots\wedge\xi_{j_k}\\&=\lambda_1\xi_1\wedge\cdots\wedge\lambda_k\xi_k=\left(\prod_{j=1}^{k}\lambda_j\right)\xi_1\wedge\cdots\wedge\xi_k,\end{aligned}$$

showing that the product is an eigenvalue. The big sum is equal to zero because in the product $\xi_1 \wedge \xi_1 \wedge \xi_{j_3} \wedge \cdots \wedge \xi_k = 0$.

When $T$ is normal we can get an orthonormal basis $\{\xi_j\}$ of eigenvectors where $T\xi_j = \lambda_j \xi_j$ (via Remark 10.6.13). By Exercise C.1.1

$$\{\xi_{j_1} \wedge \cdots \wedge \xi_{j_n} : \ j_1 < \cdots < j_n\}$$

is an orthonormal basis for $\bigwedge^k \mathcal{H}$. We have

$$\bigwedge^k T(\xi_{j_1} \wedge \cdots \wedge \xi_{j_n}) = \Big(\prod_{s=1}^{k} \lambda_{j_s}(T)\Big)\, \xi_{j_1} \wedge \cdots \wedge \xi_{j_n}.$$

Hence $\bigwedge^k T$ is diagonal with respect to this basis and so (C.10) is a list of all its eigenvalues (this was shown in Example 9.5.6, and that a diagonal operator is normal is Exercise 10.3.6). □

**C.1.1. Exercises.**

(C.1.1) Let $\mathcal{H}$ be a Hilbert space and $\{\xi_k\}$ an orthonormal basis. Show that

$$\{\xi_{j_1} \wedge \cdots \wedge \xi_{j_n} : \ j_1 < \cdots < j_n\}$$

is an orthonormal basis for $\bigwedge^n \mathcal{H}$.

(C.1.2) Let $a \in \ell^1(\mathbb{N})$. Show that

$$\sum_{j_1 < \cdots < j_k} a_{j_1} \cdots a_{j_n} \le \frac{1}{k!}\Big(\sum_{j=1}^{\infty} a_j\Big)^k.$$

(C.1.3) Let $\mathcal{H} = \mathbb{C}^n$ and $T \in \mathcal{B}(\mathcal{H}) = M_n(\mathbb{C})$. Show that $\dim \bigwedge^n \mathcal{H} = 1$, and that $\bigwedge^n T$ is the operator of multiplication by $\det T$.

(C.1.4) Let $\mathcal{H}$ be a Hilbert space and $T \in \mathcal{B}(\mathcal{H})$. Show that if $\operatorname{rank} T = n$, then $\bigwedge^{n+k}(T) = 0$ for all $k \in \mathbb{N}$.

## C.2. Lidskii's Formula

The proof of Lidskii's equality that we offer below occupies just a few lines. But if one traces the references back, particularly from Proposition C.2.4, the amount of background material required is significant. Besides the material in Section C.1, the proof also requires substantial information about compact operators (particularly, the Spectral Theorem 10.6.12), about majorization (contained in Chapter B), and a bit of non-trivial complex analysis from Section 3.11.

We begin with an estimate that looks obvious but requires the previous background for its proof.

***C.2.1. Proposition.***

*Let* $T \in \mathcal{T}(\mathcal{H})$. *Then*

$$\sum_{k=1}^{\infty} |\lambda_k(T)| \leq \mathrm{Tr}(|T|). \tag{C.12}$$

*Proof.* Fix $k \in \mathbb{N}$. As mentioned in the proof of Lemma C.1.3, we have that $|\bigwedge^k T| = \bigwedge^k |T|$. For any compact operator $S$ we always have the inequality

$$|\lambda_j(S)| \leq \|S\| = \lambda_1(|S|) \tag{C.13}$$

for any $j$. Indeed, the inequality follows by taking $\xi$ a unit eigenvector for $\lambda_j(S)$; then

$$|\lambda_j(S)| = |\langle S\xi, \xi\rangle| \leq \|S\| \, \|\xi\|^2 = \|S\|.$$

And the equality follows from

$$\|S\|^2 = \|S^*S\| = \| \, |S|^2 \, \| = \| \, |S| \, \|^2,$$

combined with Lemma 10.6.11 and Theorem 10.6.12.

As $\bigwedge^k T$ is trace-class, we can apply the above to any of its eigenvalues. Using Lemma C.1.7 on $\lambda_1(T), \dots, \lambda_k(T)$ and (C.13),

$$\prod_{j=1}^{k} |\lambda_j(T)| \leq \lambda_1(\textstyle\bigwedge^k |T|) = \displaystyle\prod_{j=1}^{k} \lambda_j(|T|).$$

The equality also follows from Lemma C.1.7, since $\bigwedge^k |T| \geq 0$ and its largest eigenvalue is obtained by multiplying the $k$ largest possible choices in (C.10), namely $\lambda_1(|T|), \dots, \lambda_k(|T|)$. By Corollary B.3.5

$$\sum_{j=1}^{k} |\lambda_j(T)| \leq \sum_{j=1}^{k} \lambda_j(|T|). \tag{C.14}$$

As this can be done for any $k$ and $\sum_{j=1}^{\infty} \lambda_j(|T|) = \mathrm{Tr}(|T|)$ (Lemma 10.7.3), the inequality (C.12) is established. $\square$

*C.2.2. Remark.* The inequalities (C.14) can be phrased very concisely in terms of majorization. If we let $\lambda(T)$ be the (infinite) vector of the eigenvalues counting multiplicities and $\mu(T)$ the vector of the singular values, then $|\lambda(T)| \prec_w \mu(T)$. In particular, $\lambda(T) \prec_w \mu(T)$.

*C.2.3. Remark.* One might be tempted to think that (C.12) could be obtained directly from $|\lambda_k(T)| \leq \lambda_k(|T|)$. This inequality holds for $k = 1$ by (C.14) but is false in general (we saw a version of this in Exercise 10.7.1). An easy example is to take $T \in M_2(\mathbb{C})$ given by

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad \text{so } |T| = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Then $\lambda_2(T) = 1$, while $\lambda_2(|T|) = \sqrt{3 - \sqrt{5}}/\sqrt{2} < 1$. Of course,

$$\lambda_1(T) + \lambda_2(T) = 2 < \sqrt{5} = \lambda_1(|T|) + \lambda_2(|T|)$$

as indicated by (C.14).

And now Next is the key technical result.

**C.2.4. Proposition.**

*Let $T \in \mathcal{T}(\mathcal{H})$ and $\{\lambda_k\}$ its list of eigenvalues, counting multiplicities. Then*

$$\det(I + zT) = \prod_{n=1}^{\infty} (1 + z\lambda_n(T)). \tag{C.15}$$

*Proof.* Suppose that $\det(I + wT) = 0$. Then $w \neq 0$ (since $\det I = 1$). By (iv) in Proposition C.1.6, $-w(-w^{-1}I - T) = I + wT$ is not invertible, so $\lambda = -w^{-1}$ is an eigenvalue for $T$. By (iii) in Proposition C.1.6, the algebraic multiplicity of $\lambda$ agrees with the order of $w$ as as zero. Conversely, if $\lambda \in \sigma(T)$ is nonzero, by (iv) in Proposition C.1.6 we have that $\det(I + wT) = 0$, so $w = -\lambda^{-1}$ is a zero of $\det(I + zT)$ with the same order as the multiplicity of $\lambda$ (again by (iii) in Proposition C.1.6).

The previous paragraph says that $z \longmapsto \det(I+zT)$ has zeros $\{z_n\}$, where $z_n = -\lambda_n(T)^{-1}$. By Proposition C.2.1 and (C.9) we have that the hypotheses in Proposition 3.11.10 are satisfied. Thus (C.15) holds. □

And now the main result follows.

**C.2.5. Theorem. (Lidskii)**

*Let $T \in \mathcal{T}(\mathcal{H})$ and $\{\lambda_n(T)\}$ its list of eigenvalues, counting multiplicities. Then*

$$\operatorname{Tr}(T) = \sum_{n=1}^{\infty} \lambda_n(T).$$

*Proof.* In (C.15), the coefficient of $z$ on the right-hand-side is $\sum_n \lambda_n(T)$; no issues with convergence nor order, as the series converges absolutely by Proposition C.2.1. And the coefficient of $z$ on the left-hand-side is $\mathrm{Tr}(T)$ by Definition C.1.4 applied to $zT$. $\square$

APPENDIX

# D The Banach–Tarski Paradox

The Banach–Tarski Paradox, which is not a paradox, is an example of how our geometric intuition does not apply to non-measurable sets. It consists of the seemingly impossible claim that a ball in $\mathbb{R}^3$ can be decomposed in a finite number of pieces such that after some rotations, said pieces can be assembled into two copies of the ball. We will first describe the construction, and afterwards briefly discuss the axiomatic setting in Section D.2.

## D.1. The Construction

**D.1.1. Definition.**

*Let $A, B$ be sets, and $\mathcal{G}$ a group that acts on each of them. We say that $A$ and $B$ are* ***equidecomposable*** *(with respect to $\mathcal{G}$), written $A \sim_e B$, if there exist $m \in \mathbb{N}$, partitions $A_1, \ldots, A_m$ of $A$ and $B_1, \ldots, B_m$ of $B$, and $g_1, \ldots, g_m \in \mathcal{G}$ such that*

$$A = \bigcup_{k=1}^{m} A_k, \qquad B = \bigcup_{k=1}^{m} B_k, \qquad g_k A_k = B_k, \ k = 1, \ldots, m.$$

The context where we will look at this will be $A, B \subset \mathbb{R}^3$, and $\mathcal{G}$ a group of rotations. In such case, equidecomposability means that we can divide $A$ into $m$ parts, rotate some or all of them, and assemble them by translation to obtain $B$.

While not entirely obvious from the definition,

***D.1.2. Lemma.***

*Equidecomposability is an equivalence relation.*

*Proof.* The non-obvious part is the transitivity. Suppose that $A \sim_e B$ and $B \sim_e C$. This means that we have disjoint unions

$$A = \bigcup_{k=1}^{m} A_k, \qquad B = \bigcup_{k=1}^{m} B_k = \bigcup_{j=1}^{n} B'_j, \qquad C = \bigcup_{j=1}^{n} C_j,$$

and $g_1, \ldots, g_m, h_1, \ldots, h_n \in \mathcal{G}$ with $g_k(A_k) = B_k$, $h_j(B'_j) = C_j$. Now define

$$A_{k,j} = g_k^{-1}(B_k \cap B'_j), \qquad C_{k,j} = h_j(B_k \cap B'_j).$$

Since each $g_k$ and $h_j$ acts injectively,

$$\bigcup_{k,j} A_{k,j} = \bigcup_{k,j} g_k^{-1}(B_k \cap B'_j) = \bigcup_{k} g_k^{-1}(B_k) = \bigcup_{k} A_k = A,$$

and similarly $\bigcup_{k,j} C_{k,j} = C$. If we define $f_{k,j} = h_j \circ g_k \in \mathcal{G}$, then

$$f_{k,j}(A_{k,j}) = h_j(B_k \cap B'_j) = C_{k,j}.$$

Hence $A \sim_e C$. □

Our goal is to prove:

***D.1.3. Theorem. (Banach–Tarski* [BT24]*)***

*There exists a group $\mathcal{G}$ of rotations in $\mathbb{R}^3$ such that the closed unit ball $B$ can be written as $B = B_1 \cup B_2$, where $B_1 \cap B_2 = \varnothing$ and $B_1 \sim_e B_2 \sim_e B$.*

In other words, the unit ball is equidecomposable with two copies of itself. How can this even be possible? It contradicts the notion of volume! Well, it really doesn't. In subsection 2.3.3, particularly Proposition 2.3.17 and Remark 2.3.18, we saw examples of subsets of $\mathbb{R}$ where the notion of length makes no sense (it is possible to have infinitely many pairwise disjoint subsets of the interval $[0, 1]$, all with "length" 1). So it is not possible to assign a notion of length to arbitrary subsets of $\mathbb{R}$, and of course the same happens with $\mathbb{R}^3$. The sets involved in the Banach–Tarski Paradox will be non-Lebesgue-measurable, and there is no contradiction; they are just too intricate for their volume to make sense.

Something that is rarely mentioned when describing this setup is that when we "assemble" the balls, we do allow the distinct parts to cross each other when translated to their place. The two balls are not "assembled" the way we assemble most physical objects, where we parts move to their final position without crossing the parts that are already in position.

Let us now work through the ideas used to prove Theorem D.1.3. The **free group** on two generators is the group $\mathbb{F}_2$ with two generators and no relations. Namely, $\mathbb{F}_2$ is the set of *reduced* words on the two generators $a, b$ and their inverses $a^{-1}, b^{-1}$. *Reduced* in this context means that any instance of $a^{-1}a, aa^{-1}, b^{-1}b, bb^{-1}$ is erased. Multiplication is defined by juxtaposition and reduction. For instance,

$$(aba^{-1}b^{-1}a)(a^{-1}bab) = ab^2.$$

The empty word is the identity, denoted $e$. The paradox is a rather straightforward consequence of the corresponding "paradox" for $\mathbb{F}_2$. Let us define, for $x \in \mathbb{F}_2$,

$$R(x) = \{y \in \mathbb{F}_2 : \ y \ \text{ begins with } x\}.$$

It is clear that

$$\mathbb{F}_2 = \{e\} \cup R(a) \cup R(a^{-1}) \cup R(b) \cup R(b^{-1}) \tag{D.1}$$

as a disjoint union. We also have

$$\mathbb{F}_2 = aR(a^{-1}) \cup R(a) = bR(b^{-1}) \cup R(b). \tag{D.2}$$

This is the "paradox" at the level of the group. If we ignore the identity point, in (D.1) we decomposed $\mathbb{F}_2$ in four disjoint parts; and (D.2) shows that by shifting two of those four components, we write $\mathbb{F}_2$ as two disjoint copies of $\mathbb{F}_2$. The Banach–Tarski Paradox consists of realizing this decomposition by using symmetries in $\mathbb{R}^3$.

Namely, we fix an irrational angle $\theta$ and we take

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \qquad B = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

That is, $A$ is a rotation around the $x$-axis, while $B$ is a rotation around the $z$-axis. With the typical choice of $\theta = \arccos\frac{1}{3}$,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & -2\sqrt{2}/3 \\ 0 & 2\sqrt{2}/3 & 1/3 \end{bmatrix}, \qquad B = \begin{bmatrix} 1/3 & -2\sqrt{2}/3 & 0 \\ 2\sqrt{2}/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We will let $\mathcal{G} = \langle A, B\rangle$ be the group generated by the two rotations $A$ and $B$. If $\mathcal{G}$ were not free, there would exist a non-trivial word $W$ on $A$ and $B$ with $W$ equal to the identity of $\mathcal{G}$. We will show that this is impossible by looking at the action of $W$ on the canonical vector $e_1$. The vectors that arise by repeated action of $A, A^{-1}, B, B^{-1}$ on $e_1$ are of a particular form, that we consider below. For $r, s, t \in \mathbb{Z}$,

$$A^{\pm 1}\begin{bmatrix} r \\ s\sqrt{2} \\ t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & \mp 2\sqrt{2}/3 \\ 0 & \pm 2\sqrt{2}/3 & 1/3 \end{bmatrix}\begin{bmatrix} r \\ s\sqrt{2} \\ t \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 3r \\ (s \mp 2t)\sqrt{2} \\ \pm 4s + t \end{bmatrix}.$$

and

$$B^{\pm 1}\begin{bmatrix} r \\ s\sqrt{2} \\ t \end{bmatrix} = \begin{bmatrix} 1/3 & \mp 2\sqrt{2}/3 & 0 \\ \pm 2\sqrt{2}/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r \\ s\sqrt{2} \\ t \end{bmatrix} = \frac{1}{3}\begin{bmatrix} r \mp 4s \\ (\pm 2r + s)\sqrt{2} \\ 3t \end{bmatrix}.$$

Let us write $\equiv$ for equivalence modulo 3. We consider four possibilities for the vector of the form $\frac{1}{3}\begin{bmatrix} r & s\sqrt{2} & t \end{bmatrix}^{\top}$:

$$1 \begin{cases} r \equiv 0 \\ s \not\equiv 0 \\ t \equiv s \end{cases} \qquad 2 \begin{cases} r + s \equiv 0 \\ s \not\equiv 0 \\ t \equiv 0 \end{cases} \qquad 3 \begin{cases} r \equiv s \\ s \not\equiv 0 \\ t \equiv 0 \end{cases} \qquad 4 \begin{cases} r \equiv 0 \\ s \not\equiv 0 \\ t + s \equiv 0 \end{cases}$$

Since $Ae_1 = e_1$, any word in $A^{\pm 1}$ and $B^{\pm 1}$ applied to $e_1$ can be considered to have a $B$ on the far right; and $Be_1$ is in case 2. We will show that, ignoring the new $\frac{1}{3}$ produced, the four operators $A, A^{-1}, B, B^{-1}$ map the cases this way:

| **A** | $\mathbf{A^{-1}}$ | **B** | $\mathbf{B^{-1}}$ |
|---|---|---|---|
| 1 $\longrightarrow$ 1 | 2 $\longrightarrow$ 4 | 1 $\longrightarrow$ 2 | 1 $\longrightarrow$ 3 |
| 2 $\longrightarrow$ 1 | 3 $\longrightarrow$ 4 | 2 $\longrightarrow$ 2 | 3 $\longrightarrow$ 3 |
| 3 $\longrightarrow$ 1 | 4 $\longrightarrow$ 4 | 4 $\longrightarrow$ 2 | 4 $\longrightarrow$ 3 |

The missing cases are not possible due to the words being reduced. Namely, since $A$ maps 1 to itself, $A^{-1}$ does not apply to it (for a word of the form $WA^{-1}$ applied to a vector $v = Az$ in 1 we would have $WA^{-1}v = WA^{-1}(Az) = Wz$ so we do a reduction instead of acting); the other three cases are mapped by $A^{-1}$ to 4, so $A$ does not apply to 4. Similarly, $B$ maps 2 to itself, so $B^{-1}$ does not apply to it; and $B^{-1}$ maps the other cases to 3, so $B$ does not apply to 3.

To check the table, note first that $Be_1$ is on 2 (recall that we are ignoring the coefficient 1/3), and:

- When we apply $A$ to any case we get $3r \equiv 0$, and $4s+t \equiv s+t \equiv s-2t$, as needed for the third condition in case 1. The only thing to check is whether in each we get that the coefficient of $\sqrt{2}$ is not a multiple of 3. In case 1, $s - 2t \equiv s - 2s = -s \not\equiv 0$. In case 2, $s - 2t \equiv s \not\equiv 0$. And in case 3, $t \equiv 0$ so $s - 2t \equiv s \not\equiv 0$.

- When we apply $A^{-1}$ to any case, we get $3r \equiv 0$, and $(s+2t)+(-4s+t) = -3s+3t \equiv 0$. This will always put us in (4) if the second entry is nonzero. When coming from (2) or (3), $s+2t \equiv s \not\equiv 0$. And when coming from (4), $s+2t = s+(s+t) \equiv s \not\equiv 0$.
- When we apply $B$ to any case we get $(r-4s)+(2r+s) = 3r-3s \equiv 0$, and $3t \equiv 0$. If coming from (1), $2r+s \equiv s \not\equiv 0$; coming from (2), $2r+s = r+(r+s) \equiv r \equiv -s \not\equiv 0$; and from case (4), $2r+s \equiv s \not\equiv 0$. So in all cases we get (2).
- When we apply $B^{-1}$ to any case we get $(r+4s)-(-2r+s) = 3r+3s \equiv 0$, $3t \equiv 0$. When coming from case (1) or (4), $-2r+s \equiv s \not\equiv 0$. And when coming from (3), $-2r+s \equiv r+s \equiv -s \not\equiv 0$.

Once we know that $We_1$ will always be in one of the four cases, we know that the second component cannot be zero, and so $We_1 \neq e_1$ for all reduced words $W$.

We have shown so far that $\mathcal{G}$ is free, with two generators, and so $\mathcal{G} \simeq \mathbb{F}_2$. Now let $S^2 \subset \mathbb{R}^3$ be the unit sphere, that is $S^2 = \{x \in \mathbb{R}^3 : \ \|x\| = 1\}$, where $\|\cdot\|$ is the Euclidean norm. We define an equivalence relation on $S^2$ by $x \sim y$ if there exists $R \in \mathcal{G}$ with $y = Rx$. Since $\mathcal{G}$ is countable and $S^2$ is not, there are uncountably many equivalence classes. Let

$$S_0 = \{x \in S : \ \exists R \in \mathcal{G}, \ Rx = x\}$$

be the set of fixed points. As each rotation fixes two points in the sphere, the set $S_0$ is countable. We will work for now on $S_1 = S^2 \setminus S_0$. Let $E$ be a set of representatives for $S_1$ (in the existence of $E$ is where we use the Axiom of Choice). By definition we have $\mathcal{G}E = S_1$, and $\mathcal{G}x \cap \mathcal{G}y = \varnothing$ for all distinct $x, y \in E$.

Recall the notation from (D.1) and (D.2). Let

$$P = \{A^{-n}x : \ n \in \mathbb{N}, \ x \in E\},$$

and

$$\begin{aligned} T_1 &= R(A)E \cup E \cup P \\ T_2 &= R(A^{-1})E \setminus P \\ T_3 &= R(B)E \\ T_4 &= R(B^{-1})E. \end{aligned}$$

From (D.1) we see that $T_1 \cup T_2 \cup T_3 \cup T_4 = S_1$. As $E$ is made up of representatives, the sets $T_j$ are pairwise disjoint. Let us see that

$$\begin{aligned} AT_2 &= T_2 \cup T_3 \cup T_4 \\ BT_4 &= T_1 \cup T_2 \cup T_4. \end{aligned}$$

For the first equality, note that $T_2 = \{A^{-n}B^{\pm1}Wx :\ x \in E,\ n \geq 1\}$. So when we apply $A$, if $n > 1$ we get $A(A^{-n}B^{\pm1}Wx) = A^{-n+1}B^{\pm1}Wx \in T_2$; and when $n = 1$ we get $A(A^{-1}B^{\pm1}Wx) = B^{\pm1}Wx$, which is in $T_3$ or $T_4$. And, conversely, those equalities show that any element of $T_2 \cup T_3 \cup T_4$ can be written as an element of $AT_2$. Similarly, for the second equality we have that $T_4 = \{B^{-n}A^{\pm1}Wx :\ x \in E,\ n \geq 1\}$. When $n = 1$, $B(B^{-1}A^{\pm1}W)x = A^{\pm1}Wx \in T_1 \cup T_2$; and when $n > 1$, $B(B^{-n}A^{\pm1}W)x = B^{-n+1}A^{\pm1}Wx \in T_4$. And any element of $T_1 \cup T_2 \cup T_4$ can be written as an element of $BT_4$. Therefore we have

$$S_1 = T_1 \cup T_2 \cup T_3 \cup T_4 = T_1 \cup AT_2 = T_3 \cup BT_4. \tag{D.3}$$

We can interpret this as saying that $S_1$ can be decomposed in four pieces, $T_1$ $T_2$, $T_3$, $T_4$, and by rotating $T_2$ and $T_4$ (in a very specific way: $T_2$ $\theta$ radians (approximately 70.53°) around the $x$-axis, and $T_4$ $\theta$ radians around the $z$-axis) we can assemble two copies of $S_1$. Namely,

$$S_1 \sim_e T_1 \cup T_2 \sim_e T_3 \cup T_4.$$

Because $S_0$ is countable, there exists a line $L$ through the origin that does not intersect $S_0$. Let $R_\theta$ be the rotation of $\theta$ radians around $L$. If $R_\theta x = R_{\theta'}x$ for some $x \in S_0$, then $\theta = \theta' \pmod{2\pi}$, since $x \notin L$. Then the set $\{\theta \in [0, 2\pi) :\ \exists n \in \mathbb{N},\ x, y \in S_0,\ y = R_\theta^n x\}$ is countable. If $\theta$ not in that set, we have $R_\theta^n S_0 \cap R_\theta^m S_0 = \varnothing$ for all $m \neq n$. Let

$$T_0 = S_0 \cup R_\theta S_0 \cup R_\theta^2 S_0 \cup \cdots$$

Now

$$S^2 \setminus S_0 = (T_0 \setminus S_0) \cup (S^2 \setminus T_0) = R_\theta T_0 \cup (S^2 \setminus T_0) \sim_e T_0 \cup (S^2 \setminus T_0) = S^2.$$

Therefore $S^2 \sim_e S^2 \setminus S_0$, and so $S^2$ is equidecomposable with two copies of itself.

Next we write $B \setminus \{0\} = \bigcup_{x \in S^2} \{tx :\ t \in (0, 1]\}$, a disjoint union. For any $T \subset S^2$ we write

$$\tilde{T} = \{tx :\ x \in T,\ t \in (0, 1]\}.$$

That is, we extend $T$ as a wedge towards the origin (these wedges we have already seen in Exercise 2.7.17). For any $W \in \mathcal{G}$ and $T \subset S^2$, we have $\widetilde{WT} = W\tilde{T}$ (this is just the fact that $Wtx = tWx$). So the decomposition

$$S^2 = T_{12} \cup T_{34} \qquad \text{with } T_{12} \sim_e T_{34} \sim_e S^2,$$

yields

$$B \setminus \{0\} = \tilde{T}_{12} \cup \tilde{T}_{34} \qquad \text{with } \tilde{T}_{12} \sim_e \tilde{T}_{34} \sim_e B \setminus \{0\}.$$

It remains to account for the origin; for this we use an idea similar to the "shift by one" that was introduced in Example 1.6.9. Let

$$C_0 = \Big\{\Big(-\frac{1}{4} + \frac{1}{4}\cos t, \frac{1}{4}\sin t, 0\Big) :\ t \in [0, 2\pi)\Big\}.$$

Then $C_0 \subset B$ is a circle of radius $\frac{1}{4}$ centered at $(-\frac{1}{4}, 0, 0)$, and $(0, 0, 0) \in C_0$. Let

$$C_1 = \Big\{\Big(-\frac{1}{4} + \frac{1}{4}\cos(n-1), \frac{1}{4}\sin(n-1), 0\Big) :\ n \in \mathbb{N}\Big\}.$$

Then $C_1$ is a countable subset of $C_0$, and $0 = (0,0,0) \in C_1$. We cannot get the same point for $n \neq m$, because $\cos n = \cos m$ and $\sin n = \sin m$ imply $m = n + 2k\pi$ for some integer $k$, which is impossible. Let $R$ be the rotation of 1 radian around the centre of $C_0$. Then $RC_1 = C_1 \setminus \{0\}$. Hence

$$B \setminus \{0\} = (C_1 \setminus \{0\}) \cup (B \setminus C_1) = RC_1 \cup (B \setminus C_1) \sim_e C_1 \cup (B \setminus C_1) = B.$$

That is, $B$ is equidecomposable with two copies of itself.

What Banach–Tarski shows is that our intuition about sets and volumes is very limited when we want to deal with arbitrary subsets of $\mathbb{R}^3$. The sets that appear in the decomposition of the ball are non-measurable; covering them with little balls in order to approximate their volume does not produce a coherent value (as it was with the Vitali sets in Proposition 2.3.17).

Our initial decomposition partitioned $S_1$ into four pieces. A quick look at the proof of the transitivity of equidecomposition shows that it is achieved by partitioning the original set into as many pieces as the product of the number of pieces required by each decomposition. In the argument we started by dividing $S_1$ into 4 pieces; the equivalence with $S^2$ required two pieces, and the equivalence between $B$ and $B \setminus \{0\}$ required another two pieces. So our construction required us to divide $B$ into $16 = 4 \times 2 \times 2$ pieces. It was proven in [**Rob47**] that the Banach–Tarski decomposition can be achieved with as little as five pieces, and that five pieces is the minimum possible.

In way more generality than the results above, Banach and Tarski proved in [**BT24**] that any two bounded subsets of $\mathbb{R}^3$ with nonempty interior are equidecomposable. This is not so far fetched as it sounds at first. If we have a ball inside our set, we could think of the set as surrounding the sphere of said ball, and repeat the above construction with little change. In particular, a very small ball can be reassembled into a very large ball, or even many of them!

## D.2. The Axiomatic Issue

The Banach–Tarski construction relies on the Axiom of Choice in an essential way (in the existence of the set $E$). The construction is so unintuitive—“we can split the ball in five parts and reassemble them as two balls of the same size as the original one”—that many feel that this is wrong, and that is why Banach–Tarski is a “paradox”. It is a big assumption, though, that arbitrary subsets of $\mathbb{R}^3$ should behave in a way that corresponds with our intuition. The existence of both the Vitali sets and Banach–Tarski challenge our intuition, but once it is clear that we are dealing with non-measurable sets, where our ideas of volume cannot possibly make sense, the paradox is not so. It is also worth mentioning that we are dealing with uncountable sets, something we cannot possibly find in reality; it just happens that in many cases the mathematical uncountability of sets does not cause problems (like intervals in Calculus) while it does in other cases. The Cantor function (page 87) was already an example of very unintuitive behaviour.

So far we have assumed the Axiom of Choice and gotten many crucial consequences out of it; thus, at least from a practical point of view, it is very useful to us. Not all mathematicians are happy with the Axiom of Choice, though. Mostly because they feel the consequences are uncomfortably non-intuitive, and also because it is true that many results can be obtained with slightly weaker axioms—in particular Countable Choice, where you only allow countably many sets in the axiom. Here are some results that are equivalent to the Axiom of Choice:

- Zorn's Lemma;
- the Well Ordering Principle;
- given two sets, their cardinalities are comparable: that is, there always exists an injection from one of them into the other;
- a function is surjective if and only if it has a right inverse;
- every vector space has a basis, and any two bases of a vector space have the same cardinality;
- every unital ring has a maximal ideal;
- if $\mathcal{X}$ is any normed space, then $\overline{B_1^{\mathcal{X}^*}(0)}$ has an extreme point;
- the Cartesian product of any family of compact sets is compact (Tychonoff's Theorem).

Next is a list of some results that require the Axiom of Choice or a slightly weaker form of it:

- Banach–Tarski Paradox;
- existence of Lebesgue non-measurable sets;
- the Hahn–Banach Theorem;
- every Hilbert space has an orthonormal basis;
- the Krein–Milman Theorem;
- the Banach–Alaoglu Theorem;
- the Baire Category Theorem (and, hence, the Open Mapping Theorem, the Closed Graph Theorem, the Inverse Mapping Theorem).

These lists do not attempt to be complete by any means. The point is that there are huge and decisive implications from the Axiom of Choice.

Let us argue now that negating the Axiom of Choice does not seem like a good position, at least from the point of view of the topics we cover in this text. For starters, negating the Axiom of Choice means assuming the existence of an empty Cartesian product of a family of nonempty sets (or, at least, saying that constructing the Cartesian product is impossible); there are mathematicians that are not bothered by this, but one has to question why Banach–Tarski is weird but this isn't? Below we list several consequences that can appear if one develops mathematics without the Axiom of Choice. Most of the list below is mentioned, with proofs, in [Jec08].

- There is a model of the ZF theory where $\mathbb{R}$ has an infinite subset without a countable subset. This means that in such model there exists $S \subset \mathbb{R}$ and $a \in \overline{S}$ such that there is no sequence $\{x_n\} \subset S$ with $x_n \to a$.
- In the same model as above, there is a subset $S \subset \mathbb{R}$ that is neither closed nor bounded, but such that every sequence in it admits a convergent subsequence.
- There exists a model of ZF where there is a vector space with no basis.
- There is a model of ZF in which there is a field with no algebraic closure.
- There exists a model of ZF—in fact, all known models with no choice—in which there is a set $S$ and an equivalence relation $\sim$ on $S$ such that there is no injection $S/\sim\, \to S$. That is, the set $S$ can be partitioned in a larger number of classes than the number of elements of the set.
- There is a model of ZF where $\mathbb{R}$ is uncountable but it can be written as a countable union of countable sets.
- There is a model of ZF where there is a vector space with two basis of different cardinalities.
- There is a model of ZF where all subsets of $\mathbb{R}^n$ are measurable (Solovay's model). In such model, though, it is possible to show that $\mathcal{P}(\mathbb{N})$ admits an equivalence relation where the quotient has larger cardinality than $\mathcal{P}(\mathbb{N})$ (note that $\mathcal{P}(\mathbb{N})$ is equipotent with $\mathbb{R}$, so this means that in this model one can partition $\mathbb{R}$ into more classes than $\mathbb{R}$ has elements). And more importantly from a measure theory point of view, Lebesgue measure is no longer $\sigma$-additive.
- There is a model of ZF where $\ell^\infty(\mathbb{N})^* = \ell^1(\mathbb{N})$. This looks fairly nice, even convenient, until one realizes that it means that the (infinite-dimensional, non-separable) Banach space $\ell^\infty(\mathbb{N})/c_0$ has trivial dual (that is, it has no continuous linear functionals; see Exercise 5.7.4).

There might be contexts where some of these consequences might be desired or tolerated, but they are certainly not desirable from the point of view of Measure Theory and Functional Analysis. The possibility of having more equivalence classes than elements seems in itself like a good reason to favour the Axiom of Choice if one cares about intuition. That said, there are branches of mathematics where negating the Axiom of Choice can make sense. If a mathematical theory does not use cardinality of sets in any sense, for instance, certain models with no choice can be useful.

APPENDIX

# E

# Ultrafilters

It is a trivial fact that not every sequence in $\ell^\infty(\mathbb{N})$ is convergent. Now, in the subspace $c_c(\mathbb{N})$ of the convergent sequences, there is a functional that clearly does not come from $\ell^1(\mathbb{N})$: the functional $f \to \lim_k f(k)$. This functional, besides being linear and bounded, is positive, and multiplicative. And the question is, does this functional admit a bounded extension to all of $\ell^\infty(\mathbb{N})$, the extension being positive and multiplicative? The (surprising) answer is: **yes, many of them**. The answer is surprising because while Hahn–Banach will provide bounded linear extensions (and these will be positive in this case), there is no guarantee that any of them will be multiplicative.

## E.1. First abstract approach: Gelfand–Naimark

Since $\ell^\infty(\mathbb{N})$ is a unital C$^*$-algebra, there is a compact set $X$ (its character space) with $\ell^\infty(\mathbb{N}) \simeq C(X)$ (Theorem 11.2.3). The topology on $X$ is the weak$^*$-topology. Now consider, for each $k \in \mathbb{N}$, the functional $\omega_k(f) = f(k)$. Since each $\omega_k$ is a character, the sequence $\{\omega_k\}$ lies inside the compact set $X$. Then there exists $\omega \in X$ and a convergent subnet $\{\omega_{k_j}\}$ such that $\omega_{k_j} \underset{j}{\to} \omega$ (weak$^*$, which is to say pointwise). This subnet has the incredible property that $\{f(k_j)\}_j$ converges for all $f \in \ell^\infty(\mathbb{N})$. Now define a functional $\omega \in \ell^\infty(\mathbb{N})^*$ by $\omega(f) = \lim_j f(k_j)$.

## E.2. Second abstract approach: the Stone Čech compactification

Note that $\ell^\infty(\mathbb{N}) = C_b(\mathbb{N})$ (with the discrete topology). By Theorem 7.4.14, there exists a compact set $\beta\mathbb{N}$, the Stone-Čech compactification of $\mathbb{N}$, such that $\mathbb{N}$ is dense in $\beta\mathbb{N}$, and every $f \in C_b(\mathbb{N})$ extends to an $\tilde{f} \in C(\beta\mathbb{N})$. Take any $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$, and define the functional $\varphi_\omega \in \ell^\infty(\mathbb{N})^*$ to be $\varphi_\omega(f) = f(\omega)$.

## E.3. A more intuitive approach: Ultrafilters

Consider first $f \in \ell^\infty(\mathbb{N})$ such that $f(k) \in \{0, 1\}$ for all $k$. Let us assume that $\varphi$ is a positive and multiplicative functional that extends lim (which we know exists by the previous abstract approaches, but in any case let us assume we have it). Being multiplicative, $\varphi(1) = 1$ (unless $\varphi$ is identically 0). Let $\alpha = \varphi(f)$. Since $f(1 - f) = 0$, we have

$$0 = \varphi(f(1 - f)) = \varphi(f)\,\varphi(1 - f) = \alpha(1 - \alpha).$$

We conclude that $\alpha \in \{0, 1\}$. Of course, we don't know which one! It might be possible to make a choice, but based on what?

Let us use the usual identification between $\mathcal{P}(\mathbb{N})$ and $\{f : \mathbb{N} \to \{0, 1\}\}$, via $A \to 1_A$. Now define a collection

$$\mathcal{U} = \{A : \ \varphi(1_A) = 1\}$$

. Then

(i) $\mathbb{N} \in \mathcal{U}$ (because $\varphi(1) = 1$);
(ii) $A \in \mathcal{U} \iff \mathbb{N} \setminus A \notin \mathcal{U}$ (because $\varphi(1 - f) = 1 - \varphi(f)$);
(iii) $A, B \in \mathcal{U} \iff A \cap B \in \mathcal{U}$ (because $1_{A\cap B} = 1_A 1_B$);
(iv) $A \in \mathcal{U}$, $A \subset B \implies B \in \mathcal{U}$ (because $\varphi$ is positive);
(v) $\{k : \ k \geq n\} \in \mathcal{U}$ for all $n \in \mathbb{N}$ (because $\varphi$ extends lim and $\lim 1_{\{1,\ldots,n\}} = 1$).

Such a $\mathcal{U}$ is called an **ultrafilter** if it satisfies (i)–(iv), and a **free ultrafilter** if it satisfies (i)–(v). The ultrafilter could also come from the multiplicative functional $f \longmapsto f(n)$ for a fixed $n$. In that case it satisfies (i)–(iv), and in fact $\mathcal{U} = \{A \subset \mathcal{P}(\mathbb{N}) : n \in A\}$; such an ultrafilter is **principal**.

The following property of ultrafilters is key:

***E.3.1. Lemma.***

*Let $\mathcal{U}$ be a free ultrafilter, and let $A_1, \ldots, A_r$ be disjoint sets with $A_1 \cup \cdots \cup A_r = \mathbb{N}$. Then there exists $j$ with $A_j \in \mathcal{U}$ and $A_i \notin \mathcal{U}$ for any $i \neq j$.*

*Proof.* We know from above that $\mathcal{U}$ corresponds to a unique linear, multiplicative and unital $\varphi$ acting on $\{f : \mathbb{N} \to \{0,1\}\}$. The set equality from above reads

$$1_{A_1} + 1_{A_2} + \cdots + 1_{A_r} = 1.$$

Applying $\varphi$, we see that one and only one of the sets can have $\varphi(A_j) = 1$. □

The proof above can also be obtained without appealing to functionals, in this way: from the properties above, $A_1 \in \mathcal{U}$ iff $A_1^c \notin \mathcal{U}$. So, if $A_1 \in \mathcal{U}$, we are done. Otherwise, $A_2 \cup \cdots A_r \in \mathcal{U}$, and we can restart the proof with one less set.

We saw above how a positive multiplicative functional on $\ell^\infty(\mathbb{N})$ that extends lim induces a free ultrafilter. Conversely, every free ultrafilter defines a linear, multiplicative, positive functional that extends lim. Indeed, we can define

$$\varphi_{\mathcal{U}}(1_A) = \begin{cases} 1, & A \in \mathcal{U} \\ 0, & A \notin \mathcal{U} \end{cases} \tag{E.1}$$

and we extend by linearity to every simple $c = \sum_{k=1}^n \beta_k 1_{A_k}$. This is well-defined (Exercise E.3.2). From Lemma E.3.1 we get that for such $c$, $\varphi(c) \in \{\beta_1, \ldots, \beta_n\}$. Therefore $|\varphi(c)| \leq \|c\|_\infty$, so $\varphi$ is bounded. As the space of simple functions is dense in $\ell^\infty(\mathbb{N})$ (Theorem 2.4.13) by Exercise 5.5.11 it has a unique bounded extension to $\ell^\infty(\mathbb{N})$ that will still be multiplicative. It remains to check that it extends lim.

Suppose that $c \in \ell^\infty(\mathbb{N})$ is convergent. We can write $c = \sum_{k=1}^\infty \beta_k 1_{\{k\}}$ and let $\beta = \lim_k \beta_k$. Because $\mathcal{U}$ is free, $\{k\} \notin \mathcal{U}$ for each $k$. Fix $\varepsilon > 0$; then there exists $k_0$ such that $|\beta_k - \beta| < \varepsilon$ for all $k \geq k_0$. So $\varphi_{\mathcal{U}}(1_{\{k\}}) = 0$ for all $k$ and

$$\begin{aligned} |\beta - \varphi_{\mathcal{U}}(c)| = \Big|\beta - \varphi_{\mathcal{U}}\Big(\sum_{k=1}^\infty \beta_k\, 1_{\{k\}}\Big)\Big| &= \Big|\beta - \varphi_{\mathcal{U}}\Big(\sum_{k=k_0}^\infty \beta_k\, 1_{\{k\}}\Big)\Big| \\ &= \Big|\varphi_{\mathcal{U}}\Big(\sum_{k=k_0}^\infty (\beta - \beta_k)\, 1_{\{k\}}\Big)\Big| \\ &\leq \sup\{|\beta - \beta_k| : k \geq k_0\} < \varepsilon. \end{aligned}$$

As this can be done for all $\varepsilon > 0$, $\varphi(c) = \beta = \lim_k \beta_k$.

**E.3.1. Limit Along an Ultrafilter.** Given $f \in \ell^\infty(\mathbb{N})$ we want to make sense, for an ultrafilter $\omega$, of the symbols $\lim_{n\to\omega} f(n)$. This will be simply the number $\varphi_\omega(f)$, where $\varphi_\omega$ is the ultrafilter constructed above. But it is actually a limit, in the following sense.

**E.3.2. Definition. (Limit along an ultrafilter)**

*Let $\omega$ be a free ultrafilter and $x \in \ell^\infty(\mathbb{N})$. We say that $\lim\limits_{n\to\omega} x_n = \alpha$ if for every open neighbourhood $V$ of $\alpha$ we have $\{n : x_n \in V\} \in \omega$.*

We leave it as exercises to show that $\lim_{n\to\omega} x_n = \varphi_\omega(x)$ (Exercise E.3.3) Together with the argument above this also shows that if $x$ is convergent then $\lim_{n\to\omega} x_n = \lim_{n\to\infty} x_n$.

**E.3.2. Exercises.**

(E.3.1) Let $A_0 \subset \mathbb{N}$. Show that $\mathcal{U} = \{A \subset \mathbb{N} : A \supset A_0\}$ is an ultrafilter. Show also that $\mathcal{U}$ is free if and only if $A_0$ is infinite.

(E.3.2) Show that $\varphi_{\mathcal{U}}$, as defined in (E.1) and extended by linearity to $\mathrm{span}\{e_k : k\}$, is well-defined.

(E.3.3) Let $x \in \ell^\infty(\mathbb{N})$. Show that $\lim\limits_{n\to\omega} x_n = \varphi_\omega(x)$.

(E.3.4) Let $n_0 \in \mathbb{N}$ and $\omega = \{A \subset \mathbb{N} : n_0 \in A\}$ the associated principal ultrafilter. Show that $\varphi_\omega(x) = x_{n_0}$.

(E.3.5) Let $x \in \ell^\infty(\mathbb{N})$ be given by $x(n) = (-1)^n$. Show that there exist free ultrafilters $\omega_1$ and $\omega_2$ such that $\varphi_{\omega_1}(x) = 1$ and $\varphi_{\omega_2}(x) = -1$.

APPENDIX

# F

# Unbounded Operators

While many pages in this book are devoted to bounded linear operators, unbounded operators do exist and they are in many cases very natural (see for instance Example 6.3.13). We also saw several examples of unbounded linear functionals, which are particular cases of unbounded operators. A priori, even very simple unbounded operators can be fairly nasty:

*F.0.1. Example.* When an unbounded linear operator is not closable, its graph can be fairly pathological. For instance let $\mathcal{X} = \mathbb{C}[x] \subset C[0,1]$, $\mathcal{Y} = \mathbb{C}[x] \subset C[2,3]$. Both $\mathcal{X}$ and $\mathcal{Y}$ are non-complete normed spaces. Let $T : \mathcal{X} \to \mathcal{Y}$ be the map $Tf = f$. This map is unbounded! (Exercise F.0.3) More importantly, we can do the following: let $(p,q) \in \mathcal{X} \times \mathcal{Y}$. Join their endpoints so that we get a continuous function $g \in C[0,3]$ with $g = p$ on $[0,1]$ and $g = q$ on $[2,3]$. Fix $\varepsilon > 0$. By Stone–Weierstrass (Theorem 7.4.20) there exists $f \in \mathbb{C}[x]$ with $\|f - g\| < \varepsilon$ on $C[0,3]$. Then $\|f - p\| < \varepsilon$ on $\mathcal{X}$ and $\|f - q\| < \varepsilon$ on $\mathcal{Y}$; thus $\|(f, Tf) - (p,q)\| < \varepsilon$ on $\mathcal{X} \times \mathcal{Y}$. That is, the graph of $T$ is dense on $\mathcal{X} \times \mathcal{Y}$.

As was mentioned in Example 6.3.13, the "good" unbounded operators are those that are closable and have dense domain.

**F.0.2. Definition.**

*Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{D} \subset \mathcal{X}$ a subspace. A linear map $T : \mathcal{D} \to \mathcal{Y}$ is an* ***unbounded operator*** *with* ***domain*** *$\mathcal{D}$, often denoted $\mathcal{D}(T)$. We say that $T$ is* ***densely defined*** *if $\mathcal{D}(T)$ is dense in $\mathcal{X}$. We say that $T$ is* ***closed*** *if its graph $\mathcal{G}(T) \subset \mathcal{X} \times \mathcal{Y}$ is closed. $T$ is* ***closable*** *if there exist a* *subspace $\tilde{\mathcal{D}} \subset \mathcal{X}$ such that $\mathcal{D}(T) \subset \tilde{\mathcal{D}}$ and $\overline{T} : \tilde{\mathcal{D}} \to \mathcal{Y}$ linear and closed, such that $\overline{T}|_{\mathcal{D}(T)} = T$.*

In principle, Definition F.0.2 says that a bounded operator $T : \mathcal{X} \to \mathcal{Y}$ is an unbounded operator with domain $\mathcal{X}$. In spite of this, we will mostly reserve the term *unbounded* for effectively unbounded operators. The most common unbounded operators tend to be densely defined.

It is routine to show an operator $T : \mathcal{D} \to \mathcal{Y}$ fails to be closable precisely when there exists nonzero $y \in \mathcal{Y}$ such that $(0, y) \in \overline{\mathcal{G}(T)}$ (Exercise F.0.5), and that $T$ is closable if and only if $\overline{\mathcal{G}(T)}$ is the graph of an operator (Exercise F.0.6). As mentioned in Examples F.0.1 and 6.3.13, non-closable operators exist (Exercise F.0.8).

### F.0.3. Proposition.

*Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces and $T : \mathcal{D}(T) \to \mathcal{Y}$ linear, where $\mathcal{D}(T) \subset \mathcal{X}$ is a subspace. Then $T$ is closed if and only if $\mathcal{D}(T)$ is complete with the norm $\|x\|_G = \|x\| + \|Tx\|$.*

*Proof.* Exercise F.0.1. □

A common instance of the situation described in Proposition F.0.3 is to consider $\mathcal{D}(T) = C^1[0,1] \subset C[0,1]$, $Tf = f'$, $\|f\|_G = \|f\|_\infty + \|f'\|_\infty$.

When our unbounded operator acts on a Hilbert space, we can think of its adjoint. The same ideas can be used to define adjoints on normed spaces, but we we will leave that as an exercise (Exercise F.0.16). The adjoint will only be unique when $T$ is densely defined, and that is why we just consider that case.

### F.0.4. Definition.

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, $\mathcal{D}(T) \subset \mathcal{H}$ a dense subspace, and $T : \mathcal{D}(T) \to \mathcal{K}$ linear. An* ***adjoint*** *for $T$ is a operator $T' : \mathcal{D}(T') \to \mathcal{H}$, such that*

$$\langle T\xi, \eta\rangle = \langle \xi, T'\eta\rangle, \qquad \xi \in \mathcal{D}(T),\ \eta \in \mathcal{D}(T'). \tag{F.1}$$

When $\eta \in \mathcal{D}(T')$, the linear functional $\gamma_\eta(\xi) = \langle T\xi, \eta\rangle$ is bounded. Indeed,

$$|\gamma_\eta(\xi)| = |\langle T\xi, \eta\rangle| = |\langle \xi, T'\eta\rangle| \le \|T'\eta\|\,\|\xi\|.$$

This means that

$$\mathcal{D}_* = \{\eta \in \mathcal{K} : \ \gamma_\eta \ \text{ is bounded}\} \tag{F.2}$$

is the largest possible domain for an adjoint of $T$. It turns out that an adjoint with domain $\mathcal{D}_*$ always exists, and is unique:

***F.0.5. Proposition.***

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, $\mathcal{D}(T) \subset \mathcal{H}$ a dense subspace, and $T : \mathcal{D}(T) \to \mathcal{K}$ linear. Let $\mathcal{D}_*$ as in (F.2). Then there exists a unique linear operator $T^* : \mathcal{D}_* \to \mathcal{H}$ that is an adjoint for $T$. We have*

$$\mathcal{G}(T^*) = V\mathcal{G}(T)^\perp, \tag{F.3}$$

*where $V$ is the isometry $V(\xi, \eta) = (\eta, -\xi)$.*

*Proof.* It is easy to check that $V$ is an isometry, and in particular this means that $V\mathcal{H}_0^\perp = (V\mathcal{H}_0)^\perp$ for any subspace $\mathcal{H}_0 \subset \mathcal{H} \times \mathcal{K}$ (Exercise F.0.2). We first show the existence of $T^*$. Let $\eta \in \mathcal{D}_*$. As $\gamma_\eta$ is bounded and $\mathcal{D}(T)$ is dense, $\gamma_\eta$ can be extended to a bounded linear functional on all of $\mathcal{H}$ (Exercise 5.5.11). By the Riesz Representation Theorem (Theorem 4.5.4) there exists a unique $T^*\eta \in \mathcal{H}$ such that

$$\gamma_\eta(\xi) = \langle \xi, T^*\eta \rangle, \qquad \xi \in \mathcal{H}.$$

The uniqueness in the Riesz Representation Theorem forces $T^*$ to be linear, and (F.1) is satisfied by construction. So $T^*$ is an adjoint for $T$.

As for the uniqueness, if $T^*$ and $T'$ are adjoints for $T$ with domain $\mathcal{D}_*$, for any $\eta \in \mathcal{D}_*$ and $\xi \in \mathcal{H}$ we have

$$\langle \xi, T'\eta \rangle = \langle T\xi, \eta \rangle = \langle \xi, T^*\eta \rangle.$$

This forces $T'\eta = T^*\eta$.

Now we prove (F.3). Let $(\eta, T^*\eta) \in \mathcal{G}(T^*)$. For any $(\xi, T\xi) \in \mathcal{G}(T)$,

$$\langle (\xi, T\xi), (T^*\eta, -\eta) \rangle = \langle \xi, T^*\eta \rangle - \langle T\xi, \eta \rangle = 0.$$

So $(T^*\eta, -\eta) \in \mathcal{G}(T)^\perp$. Then $(-\eta, -T^*\eta) \in V\mathcal{G}(T)^\perp$, and as the latter is a subspace, $(\eta, T^*\eta) \in V\mathcal{G}(T)^\perp$. Thus $\mathcal{G}(T^*) \subset V\mathcal{G}(T)^\perp$.

Conversely, suppose that $(\eta, \nu) \in V\mathcal{G}(T)^\perp$. This means that, for all $\xi \in \mathcal{D}(T)$,

$$0 = \langle (\xi, T\xi), (\nu, -\eta) \rangle = \langle \xi, \nu \rangle - \langle T\xi, \eta \rangle.$$

Then $\gamma_\eta(\xi) = \langle \xi, \nu \rangle$, and as $\mathcal{D}(T)$ is dense, it follows that $\gamma_\eta$ is bounded. Then $\eta \in \mathcal{D}(T^*)$ and $T^*\eta = \nu$. Thus $V\mathcal{G}(T)^\perp \subset \mathcal{G}(T^*)$. ☐

When $T$ is densely defined, $T^*$ is closed (Exercise F.0.7).

*F.0.6. Example.* If $T$ is as in Example F.0.1, then $\mathcal{D}_* = \{0\}$ (Exercise F.0.12).

*F.0.7. Example.* It is possible for the domain of the adjoint to be non-trivial but $T^* = 0$. Indeed, if $T$ is the map $Tf = f(0)$ defined on $\mathcal{D} = C[0,1] \subset L^2[0,1]$, then $\mathcal{D}_* \neq \{0\}$, but $T^* = 0$ (Exercise F.0.11).

It is not by chance that the examples above use unclosable operators. For we have

***F.0.8. Proposition.***

*Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, $\mathcal{D} \subset \mathcal{H}$ dense, and $T : \mathcal{D} \to \mathcal{K}$ linear. The following statements are equivalent:*

*(i) $T$ is closable;*

*(ii) $T^*$ is densely defined.*

*Proof.* (i) $\Longrightarrow$ (ii) Suppose that $\mathcal{D}(T^*)$ is not dense. Then there exists $\nu \in \mathcal{D}(T^*)^\perp$, nonzero. We have $(\nu, 0) \in \mathcal{G}(T^*)^\perp$, for

$$\langle (\nu, 0), (\eta, T^*\eta) \rangle = \langle \nu, \eta \rangle = 0.$$

Then $(0, \nu) \in V\mathcal{G}(T^*)^\perp$, which means that $V\mathcal{G}(T^*)^\perp$ is not the graph of an operator. But (using Exercise F.0.2)

$$\overline{\mathcal{G}(T)} = \mathcal{G}(T)^{\perp\perp} = (V\mathcal{G}(T^*))^\perp = V\mathcal{G}(T^*)^\perp.$$

Thus $\overline{\mathcal{G}(T)}$ is not the graph on an operator, and $T$ is not closable.

(ii) $\Longrightarrow$ (i) Suppose that $T$ is not closable. Then there exists nonzero $\nu \in \mathcal{K}$ such that $(0, \nu) \in \overline{\mathcal{G}(T)}$. This means that there is a sequence $\{\xi_n\} \subset \mathcal{D}(T)$ with $\xi_n \to 0$ and $T\xi_n \to \nu$. Then, for any $\eta \in \mathcal{D}(T^*)$,

$$\langle \nu, \eta \rangle = \lim_n \langle T\xi_n, \eta \rangle = \lim_n \gamma_\eta(\xi_n) = 0$$

since $\gamma_\eta$ is bounded. Then $\nu \in \mathcal{D}(T^*)^\perp$, which implies that $\mathcal{D}(T^*)$ cannot be dense. $\square$

*F.0.9. Example.* Let $\mathcal{X} = \mathcal{Y} = C[0,1]$, with the infinity norm. Let $\mathcal{D} = C^1[0,1]$ and $T : \mathcal{D} \to \mathcal{Y}$ the operator $Tf = f'$. Then $T$ is unbounded and closed. If instead we consider the operator $Sf = f'$ but now $\mathcal{D}(S) = C^\infty[0,1]$, this operator is closable with closure $T$ (Exercise F.0.4).

### F.0.1. Exercises.

(F.0.1) Prove Proposition F.0.3.

(F.0.2) Show that $V$, defined in Proposition F.0.5 is an isometry, and $V\mathcal{H}_0^\perp = (V\mathcal{H}_0)^\perp$ for any subspace $\mathcal{H}_0 \subset \mathcal{H} \times \mathcal{K}$.

(F.0.3) Show that the map $T$ in Example F.0.1, that is $T : \mathcal{X} \to \mathcal{Y}$ given by $Tf = f$, is unbounded.

(F.0.4) Let $\mathcal{X} = \mathcal{Y} = C[0,1]$, with the infinity norm. Let $\mathcal{D} = C^1[0,1]$ and $T : \mathcal{D} \to \mathcal{Y}$ the operator $Tf = f'$. Show that $T$ is unbounded and closed. If instead we consider the operator $Sf = f'$ but now $\mathcal{D}(S) = C^\infty[0,1]$, show that this operator is closable with closure $T$.

(F.0.5) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{D} \subset \mathcal{X}$ a subspace, and $T : \mathcal{D} \to \mathcal{Y}$ linear. Show that the following statements are equivalent:

(i) $T$ is closable;

(ii) $(0, y) \in \overline{\mathcal{G}(T)}$ implies $y = 0$.

(F.0.6) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{D} \subset \mathcal{X}$ a subspace, and $T : \mathcal{D} \to \mathcal{Y}$ linear. Show that $T$ is closable if and only if $\overline{\mathcal{G}(T)}$ is the graph of an operator.

(F.0.7) Let $\mathcal{D}(T) \subset \mathcal{H}$ be dense, and $T : \mathcal{D}(T) \to \mathcal{K}$ linear. Show that $T^*$ is closed.

(F.0.8) Let $\mathcal{X} = L^2[0,1]$ with its dense subspace $\mathcal{D} = C[0,1]$. Define $T : \mathcal{D} \to \mathcal{D}$ by $Tf = f(0)$. Show that this operator is unbounded. Also, consider the functions $f_n = (1 - nt)\, 1_{[0,\frac{1}{n}]}$ and show that $f_n \in C[0,1]$ for all $n$, $\|f_n\|_2 \to 0$, and $Tf_n = 1$ for all $n$, so $\mathcal{G}(T)$ is not closed; conclude that $T$ is not closed and that it is not even closable.

(F.0.9) Give an example of a Banach space $\mathcal{X}$, subspaces $\mathcal{D}, M \subset \mathcal{X}$, and an idempotent $E : \mathcal{D} \to M$ which is unbounded.

(F.0.10) Let $p \in [1,\infty)$, $\mathcal{X} = \mathcal{Y} = \ell^p(\mathbb{N})$, $\mathcal{D} = c_{00}$, and $T : \mathcal{D} \to \mathcal{X}$ given by

$$T\Big(\sum_{k=1}^n c_k e_k\Big) = \Big(\sum_{k=1}^n k\, c_k\Big)\, e_1 + \sum_{k=2}^n c_k e_k.$$

Decide if $T$ is closable. Find $T^*$.

(F.0.11) Let $T : \mathcal{D} \to C[0,1]$ be the map $Tf = f(0)$, where $\mathcal{D} = C[0,1] \subset L^2[0,1]$. Show that $\mathcal{D}_* \neq \{0\}$, but $T^* = 0$.

(F.0.12) Let $T$ be as in Example F.0.1. Show that $\mathcal{D}_* = \{0\}$.

(F.0.13) Let $\mathcal{H} = \mathcal{K} = \ell^2(\mathbb{N})$, $\mathcal{D} = c_{00}$, and $T : \mathcal{D} \to \mathcal{K}$ the linear operator induced by $Te_{p^n} = e_p$ for each $p \in \mathbb{N}$ prime and $n \in \mathbb{N}$. Find $\mathcal{D}(T^*)$.

(F.0.14) Let $\mathcal{D} \subset \mathcal{H}$ be dense and $T : \mathcal{D} \to \mathcal{K}$ linear. Show that the following statements are equivalent:

(i) $\mathcal{D}(T^*) = \{0\}$;

(ii) $\overline{\mathcal{G}(T)} = \mathcal{H} \times \mathcal{K}$.

(F.0.15) Let $g \in L^{\infty}(\mathbb{R})$, and such that $\int_{\mathbb{R}} |g|^2 = \infty$. Fix $h_0 \in L^2(\mathbb{R})$. Let $T : \mathcal{D}(T) \to L^2(\mathbb{R})$ be given by

$$Tf = \langle f, g \rangle h_0,$$

where

$$\mathcal{D}(T) = \left\{ f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} |fg| < \infty \right\}.$$

Show that $T$ is densely defined, and find $T^*$.

(F.0.16) For an unbounded operator $T : \mathcal{D}(T) \to \mathcal{Y}$, where $\mathcal{D}(T) \subset \mathcal{X}$ is dense and $\mathcal{X}, \mathcal{Y}$ are normed spaces, write a definition for $T^*$, and explore how much of the results in the text can be made to work. Reflexivity might be needed in some cases.

(F.0.17) Continuing from Exercise F.0.16, show that if $T$ is closable and $\mathcal{X}, \mathcal{Y}$ are reflexive, then

$$\overline{T} = T^{**}.$$

This requires showing first that because $T$ is closable then $T^*$ is densely defined, which hopefully was done in Exercise F.0.16.

(F.0.18) Let $\mathcal{X} = \mathcal{Y} = c_0$, and $T : \mathcal{X} \to \mathcal{Y}$ given by

$$Tx = (x_1, 2x_2, 3x_3, \ldots),$$

with $\mathcal{D}(T) = c_{00}$. Decide if $T$ is closed or not, and find $T^*$.

(F.0.19) Let $\mathcal{X}, \mathcal{Y}$ be normed spaces and $T : \mathcal{D}(T) \to \mathcal{Y}$ linear and densely defined. Show that $\mathcal{D}(T^*) = \mathcal{Y}^*$ if and only if $T$ is bounded.

# Bibliography


[AAP86] C. A. Akemann, J. Anderson, and G. K. Pedersen, *Excising states of $C^*$-algebras*, Canad. J. Math **38** (1986), no. 5, 223–230.

[AGR23] R. J. Archbold, I. Gogić, and L. Robert, *Local variants of the Dixmier property and weak centrality for* $\mathrm{C}^*$*-algebras*, Int. Math. Res. Not. IMRN (2023), no. 2, 1483–1513.

[AH11] S. A. Argyros and R. G. Haydon, *A hereditarily indecomposable $L_\infty$-space that solves the scalar-plus-compact problem*, Acta Math. **206** (2011), no. 1, 1–54.

[AK06] F. Albiac and N. J. Kalton, *Topics in banach space theory*, vol. 233, Springer, 2006.

[AM08] M. Argerami and P. Massey, *The local form of doubly stochastic maps and joint majorization in* $\mathrm{II}_1$ *factors*, Integral Equations Operator Theory **61** (2008), no. 1, 1–19.

[Apé79] R. Apéry, *Irrationalité de $\zeta 2$ et $\zeta 3$*, no. 61, 1979, Luminy Conference on Arithmetic, pp. 11–13.

[APT73] C. A. Akemann, G. K. Pedersen, and J. Tomiyama, *Multipliers of $c^*$-algebras*, J. Funct. Anal. **13** (1973), 277–301.

[Arc77] R. J. Archbold, *An averaging process for $c^*$-algebras related to weighted shifts*, Proc. London Math. Soc. (3) **35** (1977), 541–554.

[Arc79] ———, *On the dixmier property of certain algebras*, Math. Proc. Camb. Phil. Soc. **86** (1979), 251–259.

[ART17] R. Archbold, L. Robert, and A. Tikuisis, *The Dixmier property and tracial states for $C^*$-algebras*, J. Funct. Anal. **273** (2017), no. 8, 2655–2718.

[Arv69] W. Arveson, *Subalgebras of $C^*$-algebras*, Acta Math. **123** (1969), 141–224.

[Arv74] ———, *On groups of automorphisms of operator algebras*, J. Functional Analysis **15** (1974), 217–243.

[AW69] H. Araki and E. J. Woods, *A classification of factors*, Publ. Res. Inst. Math. Sci. Ser. A **4** (1968/69), 51–130.

[Ber72] S. K. Berberian, *Baer *-rings*, Springer-Verlag, New York, 1972, Die Grundlehren der mathematischen Wissenschaften, Band 195.

[BES94] H. Bart, T. Ehrhardt, and B. Silbermann, *Zero sums of idempotents in Banach algebras*, Integral Equations Operator Theory **19** (1994), no. 2, 125–134.

[Bha97] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York, 1997.

[Bir46] G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Univ. Nac. Tucumán. Revista A. **5** (1946), 147–151.

[BO08] N. P. Brown and N. Ozawa, *$C^*$-algebras and finite-dimensional approximations*, Graduate Studies in Mathematics, vol. 88, American Mathematical Society, Providence, RI, 2008.

[BT24] S. Banach and A. Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fund. math **6** (1924), no. 1, 244–277.

[CE76] M. D. Choi and E. G. Effros, *The completely positive lifting problem for $C^*$-algebras*, Ann. of Math. (2) **104** (1976), no. 3, 585–609.

[Cho79] M. D. Choi, *A simple $c^*$-algebra generated by two finite order unitaries*, Canad. J. Math. **31** (1979), 867–880.

[CI10] I. Chifan and A. Ioana, *Ergodic subequivalence relations induced by a Bernoulli action*, Geom. Funct. Anal. **20** (2010), no. 1, 53–67.

[Coh64] H. B. Cohen, *Injective envelopes of Banach spaces*, Bull. Amer. Math. Soc. **70** (1964), 723–726.

[Con70] J. B. Conway, *The numerical range and a certain convex set in an infinite factor*, J. Funct. Anal. **5** (1970), 428–435.

[Con71] A. Connes, *Un nouvel invariant pour les algèbres de von Neumann*, C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A900–A903.

[Con73] ———, *Une classification des facteurs de type* III, Ann. Sci. École Norm. Sup. (4) **6** (1973), 133–252.

[Con76] A. Connes, *Classification of injective factors*, Ann. of Math. (2) **104** (1976), no. 1, 73–115.

[Con78] J. B. Conway, *Functions of one complex variable*, second ed., Graduate Texts in Mathematics, vol. 11, Springer-Verlag, New York-Berlin, 1978.

[Con90] ———, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990.

[Con12] J. B. Conway, *Functions of one complex variable ii*, vol. 159, Springer Science & Business Media, 2012.

[Dav73] A. M. Davie, *The approximation problem for banach spaces*, Bulletin of the London Mathematical Society **5** (1973), no. 3, 261–266.

[Dav96] K. R. Davidson, *$C^*$-algebras by example*, Fields Institute Monographs, vol. 6, American Mathematical Society, Providence, RI, 1996.

[Dav25] ———, *Functional analysis and operator algebras*, CMS/CAIMS Books in Mathematics, vol. 13, Springer, Cham, [2025] ©2025.

[Dix48] J. Dixmier, *Position relative de deux variétés linéaires fermées dans un espace de Hilbert*, Revue Sci. **86** (1948), 387–399.

[Dix49] J. Dixmier, *Les anneaux d'opérateurs de classe finie*, Ann. Sci. École Norm. Sup. (3) **66** (1949), 209–261.

[Dix51] ———, *Sur certains espaces considérés par M. H. Stone*, Summa Brasil. Math. **2** (1951), 151–182.

[dlHS94] P. de la Harpe and G. Skandalis, *Les réseaux dans les groupes semi-simples ne sont pas intérieurement moyennables*, Enseign. Math. (2) **40** (1994), 291–311.

[DR50a] A. Dvoretzky and C. A. Rogers, *Absolute and unconditional convergence in normed linear spaces*, Proc. Nat. Acad. Sci. U.S.A. **36** (1950), 192–197.

[DR50b] A. Dvoretzky and C. A. Rogers, *Absolute and unconditional convergence in normed linear spaces*, Proceedings of the National Academy of Sciences **36** (1950), no. 3, 192–197.

[Dri20] D. Drimbe, *Prime* $\mathrm{II}_1$ *factors arising from actions of product groups*, J. Funct. Anal. **278** (2020), no. 5, 108366, 23.

[DS88] N. Dunford and J. T. Schwartz, *Linear operators. Part I*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[Dug66] J. Dugundji, *Topology*, Allyn and Bacon Inc., Boston, Mass., 1966.

[Dyk94] K. Dykema, *Interpolated free group factors*, Pacific J. Math. **163** (1994), no. 1, 123–135.

[Enf73] P. Enflo, *A counterexample to the approximation problem in banach spaces*, Acta mathematica **130** (1973), 309–317.

[FHH+11] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach space theory: the basis for linear and nonlinear analysis*, Springer, 2011.

[FM77a] J. Feldman and C. C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras. I*, Trans. Amer. Math. Soc. **234** (1977), no. 2, 289–324.

[FM77b] ———, *Ergodic equivalence relations, cohomology, and von Neumann algebras. II*, Trans. Amer. Math. Soc. **234** (1977), no. 2, 325–359.

[Fug50] B. Fuglede, *A commutativity theorem for normal operators*, Proc. Nat. Acad. Sci. U.S.A. **36** (1950), 35–40.

[Gab02] D. Gaboriau, *Invariants $l^2$ de relations d'équivalence et de groupes*, Publ. Math. Inst. Hautes Études Sci. (2002), no. 95, 93–150.

[GJ02] M. I. Garrido and J. A. Jaramillo, *Variations on the Banach-Stone theorem*, Extracta Math. **17** (2002), no. 3, 351–383, IV Course on Banach Spaces and Operators (Spanish) (Laredo, 2001).

[Gle58] A. M. Gleason, *Projective topological spaces*, Illinois J. Math. **2** (1958), 482–489.

[Gol56] M. Goldman, *Structure of aw*-algebras, i*, Duke Math. J. **23** (1956), 23–34.

[Gol22] I. Goldbring, *The Connes embedding problem: a guided tour*, Bull. Amer. Math. Soc. (N.S.) **59** (2022), no. 4, 503–560.

[Goo50] D. B. Goodner, *Projections in normed linear spaces*, Trans. Amer. Math. Soc. **69** (1950), 89–108.

[Gro55] A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc. **16** (1955), Chapter 1: 196 pp.; Chapter 2: 140.

[Haa15] U. Haagerup, *A new look at c*-simplicity and the unique trace property of a group*, 2015, arXiv:1509.05880v1, 8 pp.

[Hal69] P. R. Halmos, *Two subspaces*, Transactions of the American Mathematical Society **144** (1969), 381–389.

[Ham79a] M. Hamana, *Injective envelopes of $C^*$-algebras*, J. Math. Soc. Japan **31** (1979), no. 1, 181–197.

[Ham79b] ———, *Injective envelopes of operator systems*, Publ. Res. Inst. Math. Sci. **15** (1979), no. 3, 773–785.

[Has58] M. Hasumi, *The extension property of complex Banach spaces*, Tohoku Math. J. (2) **10** (1958), 135–142.

[HJ94] R. Horn and R. Johnson, Charles, *Topics in matrix analysis*, Cambridge University Press, 1994.

[HKW86] H. Halpern, V. Kaftal, and G. Weiss, *The relative dixmier property in discrete crossed products*, J. Funct. Anal. **69** (1986), 121–140.

[HN91] F. Hiai and Y. Nakamura, *Closed convex hulls of unitary orbits in von neumann algebras*, Trans. Amer. Math. Soc. **323** (1991), 1–38.

[Hof62] K. Hoffman, *Banach spaces of analytic functions*, Prentice-Hail, Inc., Englewood Cliffs, N. J., 1962.

[IPP08] A. Ioana, J. Peterson, and S. Popa, *Amalgamated free products of weakly rigid factors and calculation of their symmetry groups*, Acta Math. **200** (2008), no. 1, 85–153.

[IPV13] A. Ioana, S. Popa, and S. Vaes, *A class of superrigid group von Neumann algebras*, Ann. of Math. (2) **178** (2013), no. 1, 231–286.

[Jam51] R. C. James, *A non-reflexive Banach space isometric with its second conjugate space*, Proc. Nat. Acad. Sci. U.S.A. **37** (1951), 174–177.

[Jec08] T. J. Jech, *The axiom of choice*, Courier Corporation, 2008.

[Jon83] V. F. R. Jones, *Index for subfactors*, Invent. Math. **72** (1983), no. 1, 1–25.

[JP95] M. Junge and G. Pisier, *Bilinear forms on exact operator spaces and $B(H)\otimes B(H)$*, Geom. Funct. Anal. **5** (1995), no. 2, 329–363.

[Kad56] R. V. Kadison, *Operator algebras with a faithful weakly-closed representation*, Ann. of Math. (2) **64** (1956), 175–181.

[Kad61] ———, *The trace in finite operator algebras*, Proc. Amer. Math. Soc. **12** (1961), 973–977.

[Kap51a] I. Kaplansky, *Projections in Banach algebras*, Ann. of Math. (2) **53** (1951), 235–249.

[Kap51b] ———, *A theorem on rings of operators*, Pacific J. Math. **1** (1951), 227–232.

[Kau66] R. Kaufman, *A type of extension of Banach spaces*, Acta Sci. Math. (Szeged) **27** (1966), 163–166.

[Kel52] J. L. Kelley, *Banach spaces with the extension property*, Trans. Amer. Math. Soc. **72** (1952), 323–326.

[Ken20] M. Kennedy, *An intrinsic characterization of $C^*$-simplicity*, Ann. Sci. Éc. Norm. Supér. (4) **53** (2020), no. 5, 1105–1119.

[Kir77] E. Kirchberg, *$C^*$-nuclearity implies CPAP*, Math. Nachr. **76** (1977), 203–212.

[Kir93] ———, *On nonsemisplit extensions, tensor products and exactness of group $C^*$-algebras*, Invent. Math. **112** (1993), no. 3, 449–489.

[KLR67] R. V. Kadison, E. C. Lance, and J. R. Ringrose, *Derivations and automorphisms of operator algebras, ii*, J. Funct. Anal. **1** (1967), 204–221.

[KR67] R. V. Kadison and J. R. Ringrose, *Derivations and automorphisms of operator algebras*, Comm. Math. Phys. **4** (1967), 32–63.

[KR97a] R. V. Kadison and J. R. Ringrose, *Fundamentals of the theory of operator algebras. Vol. I*, Graduate Studies in Mathematics, vol. 15, American Mathematical Society, Providence, RI, 1997, Elementary theory, Reprint of the 1983 original.

[KR97b] ———, *Fundamentals of the theory of operator algebras. Vol. II*, Graduate Studies in Mathematics, vol. 16, American Mathematical Society, Providence, RI, 1997, Advanced theory, Corrected reprint of the 1986 original.

[LV93] Y. I. Lyubich and L. N. Vaserstein, *Isometric embeddings between classical Banach spaces, cubature formulas, and spherical designs*, Geom. Dedicata **47** (1993), no. 3, 327–362.

[Mag08] B. Magajna, *On weakly central c*-algebras*, J. Math. Anal. Appl. **342** (2008), 1481–1484.

[McD69] D. McDuff, *Uncountably many* $\mathrm{II}_1$ *factors*, Ann. of Math. (2) **90** (1969), 372–377.

[Mer85] R. Mercer, *Convergence of Fourier series in discrete crossed products of von Neumann algebras*, Proc. Amer. Math. Soc. **94** (1985), no. 2, 254–258.

[MOA11] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of majorization and its applications*, second ed., Springer Series in Statistics, Springer, New York, 2011.

[Mun00] J. R. Munkres, *Topology*, Prentice hall Upper Saddle River, 2000.

[MvN43] F. J. Murray and J. von Neumann, *On rings of operators. IV*, Ann. of Math. (2) **44** (1943), 716–808.

[Nac49] L. Nachbin, *On the Han-Banach theorem*, An. Acad. Brasil. Ci. **21** (1949), 151–154.

[Nic12] B. Nica, *The Mazur-Ulam theorem*, Expo. Math. **30** (2012), no. 4, 397–398.

[NRS16] P. W. Ng, L. Robert, and P. Skoufranis, *Majorization in c*-algebras*, 2016, Trans. Amer. Math. Soc., to appear; arXiv:1608.04350v1, 32 pp.

[NS16] P. W. Ng and P. Skoufranis, *Closed convex hulls of unitary orbits in certain simple real rank zero c*-algebra*, 2016, Canad. J. Math., to appear; arXiv:1603.07059v1, 35 pp.

[OP10] N. Ozawa and S. Popa, *On a class of* $\mathrm{II}_1$ *factors with at most one Cartan subalgebra*, Ann. of Math. (2) **172** (2010), no. 1, 713–749.

[OP16] N. Ozawa and G. Pisier, *A continuum of* $\mathrm{C}^*$*-norms on* $\mathbb{B}(H) \otimes \mathbb{B}(H)$ *and related tensor products*, Glasg. Math. J. **58** (2016), no. 2, 433–443.

[Oza13] N. Ozawa, *Dixmier approximation and symmetric amenability for c*-algebras*, J. Math. Soc. Univ. Tokyo **20** (2013), 349–374.

[Pau02] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge Studies in Advanced Mathematics, vol. 78, Cambridge University Press, Cambridge, 2002.

[Pec65] N. Peck, *On non locally convex spaces*, Mathematische Annalen **161** (1965), no. 2, 102–115.

[Ped79] G. K. Pedersen, *$C^*$-algebras and their automorphism groups*, London Mathematical Society Monographs, vol. 14, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1979.

[Pel60] A. Pelczyński, *Projections in certain banach spaces*, Studia Mathematica **2** (1960), no. 19, 209–228.

[Pel64] ———, *A proof of Eberlein-Šmulian theorem by an application of basic sequences*, Bulletin De L'Academie Polonaise Des Sciences-Serie Des Sciences Mathematiques Astronomiques Et Physiques **12** (1964), no. 9, 543.

[Pop99] S. Popa, *The relative dixmier property for inclusions of von neumann algebras of finite index*, Ann. Sci. Éc. Norm. Sup. **32** (1999), 743–767.

[Pop00] ———, *On the relative dixmier property for inclusions of c*-algebras*, J. Funct. Anal. **171** (2000), 139–154.

[Pop06a] S. Popa, *On a class of type* $\mathrm{II}_1$ *factors with Betti numbers invariants*, Ann. of Math. (2) **163** (2006), no. 3, 809–899.

[Pop06b] ———, *Strong rigidity of* $\mathrm{II}_1$ *factors arising from malleable actions of w-rigid groups. I*, Invent. Math. **165** (2006), no. 2, 369–408.

[Pop07] ———, *Cocycle and orbit equivalence superrigidity for malleable actions of w-rigid groups*, Invent. Math. **170** (2007), no. 2, 243–295.

[Pow67] R. T. Powers, *Representations of uniformly hyperfinite algebras and their associated von Neumann rings*, Ann. of Math. (2) **86** (1967), 138–171.

[Pow75] ———, *Simplicity of the C*-algebra associated with the free group on two generators*, Duke Math. J. **42** (1975), 151–156.

[PP86] M. Pimsner and S. Popa, *Entropy and index for subfactors*, Ann. Sci. École Norm. Sup. (4) **19** (1986), no. 1, 57–106.

[PS79] W. L. Paschke and N. Salinas, *$C^*$-algebras associated with free products of groups*, Pacific J. Math. **82** (1979), no. 1, 211–221.

[Put51] C. R. Putnam, *On normal operators in Hilbert space*, Amer. J. Math. **73** (1951), 357–362.

[PV10] S. Popa and S. Vaes, *Group measure space decomposition of* $\mathrm{II}_1$ *factors and $W^*$-superrigidity*, Invent. Math. **182** (2010), no. 2, 371–417.

[PV14] ———, *Unique Cartan decomposition for* $\mathrm{II}_1$ *factors arising from arbitrary actions of free groups*, Acta Math. **212** (2014), no. 1, 141–198.

[Rie82] N. Riedel, *On the dixmier property of simple c*-algebras*, Math. Proc. Camb. Phil. Soc. **91** (1982), 75–78.

[Rie83] ———, *The weak dixmier property implies the dixmier property*, Dilation theory, Toeplitz operators and other topics, vol. 11, 1983, pp. 299–301.

[Rin84] J. R. Ringrose, *On the dixmier approximation theorem*, Proc. London Math. Soc. (3) **49** (1984), 37–57.

[Rob47] R. M. Robinson, *On the decomposition of spheres*, Fund. Math **34** (1947), 246–260.

[Ră94] F. Rădulescu, *Random matrices, amalgamated free products and subfactors of the von Neumann algebra of a free group, of noninteger index*, Invent. Math. **115** (1994), no. 2, 347–389.

[Rud87] W. Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987.

[Sak70] S. Sakai, *An uncountable number of* $\mathrm{II}_1$ *and* $\mathrm{II}_\infty$ *factors*, J. Functional Analysis **5** (1970), 236–246.

[Sim05] B. Simon, *Trace ideals and their applications*, no. 120, American Mathematical Soc., 2005.

[Sko16] P. Skoufranis, *Closed convex hulls of unitary orbits in c*-algebras of real rank zero*, J. Funct. Anal. **270** (2016), 1319–1360.

[Sob41] A. Sobczyk, *Projection of the space* $(m)$ *on its subspace* $(c_0)$, Bull. Amer. Math. Soc. **47** (1941), 938–947.

[Sok11] A. D. Sokal, *A really simple elementary proof of the uniform boundedness theorem*, Amer. Math. Monthly **118** (2011), no. 5, 450–452.

[SS08] A. M. Sinclair and R. R. Smith, *Finite von Neumann algebras and masas*, London Mathematical Society Lecture Note Series, vol. 351, Cambridge University Press, Cambridge, 2008.

[Ste93] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[Sto37] M. H. Stone, *Applications of the theory of Boolean rings to general topology*, Trans. Amer. Math. Soc. **41** (1937), no. 3, 375–481.

[Sto48] ———, *The generalized Weierstrass approximation theorem*, Math. Mag. **21** (1948), 167–184, 237–254.

[SZ73] S. Strătilă and L. Zsidó, *An algebraic reduction theory for w*-algebras. ii*, Rev. Roumaine Math. Pures Appl. **18** (1973), 407–460.

[Sza78] A. Szankowski, *Subspaces without the approximation property*, Israel Journal of Mathematics **30** (1978), no. 1, 123–129.

[Tak64] M. Takesaki, *On the cross-norm of the direct product of* $C^*$*-algebras*, Tohoku Math. J. (2) **16** (1964), 111–122.

[Tak70] M. Takesaki, *Tomita's theory of modular Hilbert algebras and its applications*, Lecture Notes in Mathematics, Vol. 128, Springer-Verlag, Berlin-New York, 1970.

[Tak72] M. Takesaki, *Conditional expectations in von Neumann algebras*, J. Functional Analysis **9** (1972), 306–321.

[Tak02] M. Takesaki, *Theory of operator algebras. I*, Encyclopaedia of Mathematical Sciences, vol. 124, Springer-Verlag, Berlin, 2002, Reprint of the first (1979) edition, Operator Algebras and Non-commutative Geometry, 5.

[Tak03] ———, *Theory of operator algebras. II*, Encyclopaedia of Mathematical Sciences, vol. 125, Springer-Verlag, Berlin, 2003, , Operator Algebras and Non-commutative Geometry, 6.

[VdPA79] A. Van der Poorten and R. Apéry, *A proof that euler missed...*, The Mathematical Intelligencer **1** (1979), no. 4, 195–203.

[Vee71] W. A. Veech, *Short proof of Sobczyk's theorem*, Proc. Amer. Math. Soc. **28** (1971), 627–628.

[vN40] J. von Neumann, *On rings of operators. III*, Ann. of Math. (2) **41** (1940), 94–161.

[vN49] J. von Neumann, *On rings of operators. Reduction theory*, Ann. of Math. (2) **50** (1949), 401–485.

[Was90] S. Wassermann, *Tensor products of free-group* $C^*$*-algebras*, Bull. London Math. Soc. **22** (1990), no. 4, 375–380.

[Whi67] R. Whitley, *An elementary proof of the Eberlein-Šmulian theorem*, Mathematische Annalen **172** (1967), no. 2, 116–118.

[Wri76] J. D. M. Wright, *Wild AW*-factors and Kaplansky-Rickart algebras*, J. London Math. Soc. (2) **13** (1976), no. 1, 83–89.

[Yea71] F. J. Yeadon, *A new proof of the existence of a trace in a finite von Neumann algebra*, Bull. Amer. Math. Soc. **77** (1971), 257–260.

[Zsi74] L. Zsido, *Note on dixmier's trace type sets in properly infinite w*-algebras*, Rev. Roumaine Math. Pures Appl. **19** (1974), 269–274.

# Index

# List of Symbols